1061
3
oy 1

THE TRANSFER EFFECTS OF PRACTICE IN CANCELLATION TESTS

BY

MELVIN ALBERT MARTIN, A. M.

REPRINT OF

ARCHIVES OF PSYCHOLOGY

NO. 32

SUBMITTED IN PARTIAL FULFILMENT OF THE
REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
IN THE FACULTY OF PHILOSOPHY,
COLUMBIA UNIVERSITY

NEW YORK
1915

THE TRANSFER EFFECTS OF PRACTICE IN CANCELLATION TESTS

BY

MELVIN ALBERT MARTIN, A. M.

REPRINT OF

ARCHIVES OF PSYCHOLOGY

NO. 32

SUBMITTED IN PARTIAL FULFILMENT OF THE
REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
IN THE FACULTY OF PHILOSOPHY,
COLUMBIA UNIVERSITY

NEW YORK

1915

LB1061
.M3

GIft
The University
MAR 4 1915

TABLE OF CONTENTS

Introduction

CHAPTER I. Discussion of Previous Investigations. 1

CHAPTER II. Description and Discussion of Materials and Procedure..... 23

Section 1. Preliminary Investigations..... 23
A. Reaction Time.
B. Cancellation.
C. Cancellation.

Section 2. The Present Investigation..... 24
A. Practice Material.
B. Test Material.
C. Procedure.

Section 3. Treatment of Data..... 28
A. Data of the Practice Series.
B. Data of the Test Series.

CHAPTER III. Interpretation of Results..... 46

Section 1. In Terms of Speed and Accuracy of Performance 47

Section 2. When Corrected for Errors..... 62

Section 3. Summary of Results..... 65

Section 4. General Conclusions. 67

INTRODUCTION

The problem of Formal Discipline, or the transfer effects of training, or the spread of special training, is no longer the question whether the special training of a function or a function-group increases its efficiency with all sorts of materials and in all kinds of situations, but it is, rather, under what conditions, to what extent, and in what direction, may we expect training of a special kind to produce transfer effects. In planning the present investigation the author had these questions in mind. No one realizes more than he the difficulties involved in conducting and completing such an experiment under conditions that make for scientific accuracy. Such a research requires quite a large number of subjects divided into two groups—one for the control, to determine the improvement in the tests themselves, and one for practice. To secure a sufficient number of subjects for the control is comparatively easy, but to secure a suitable practice group and hold them to systematic and prolonged practice under conditions that favor a maximum of measurable improvement from practice period to practice period is not an easy task.

The investigation is purposely limited to a narrow field in order to avoid possible sources of error due to complicating the procedure as well as the results by attempting too much in one experiment. It is a straightforward attempt to discover the transfer effects of prolonged practice in cancelling in one situation upon cancelling in other situations. The materials in both the practice series and the test series, as well as the methods of procedure were such as to insure exact quantitative scoring; and both groups of subjects were sufficiently large to prevent chance errors and individual variations from invalidating the results.

Chapter I contains a survey and criticism of previous investigations, with suggestions concerning the necessity and importance of standardizing procedure in this field of research. In chapter II the materials and procedure used in the present investigation are presented and discussed. This chapter con-

tains also the data of both the practice and the tests, with a brief discussion of the same. Chapter III deals with the transfer effects. The results of the experiment are interpreted and explained and general conclusions drawn.

The writer gratefully acknowledges assistance from several sources. He is indebted to Professors R. S. Woodworth, J. McK. Cattell, E. L. Thorndike, and Dr. A. T. Poffenberger, Jr., for many helpful suggestions and criticisms; and to the Hebrew Orphans Home, of New York City, for the services of the subjects and for the use of the school rooms of that institution. To the subjects themselves he is indebted for unfailing devotion to the work involved, especially in the practice.

THE TRANSFER EFFECTS OF PRACTICE IN CANCELLATION TESTS

CHAPTER I.

DISCUSSION OF PREVIOUS INVESTIGATIONS

Ever since James announced the results of his experiment on the transfer effects of memory-training, and Thorndike and Woodworth published their conclusions based on the results of an elaborate series of transfer experiments, this field of research has proved a fruitful source of experimentation. While it is true that James found no evidence of transfer, yet, it is to James we must make acknowledgement for the inception of investigations which have already modified our psychological and educational concepts, and which bid fair to revolutionize our educational procedure.

However much investigators may disagree as to whether special training is general in its effects, they are agreed that there is no function which through special training is made equally effective in all sorts of situations. It is hardly necessary to mention that conclusions concerning transfer based on experiments on cross-education should be rejected, for Professor Thorndike has already pointed out that "they have been improperly used as evidence on our question."[1]

In James's[2] experiment on the transfer effects of memory training there was not only no evidence of transfer, but not even of improvement in the tests themselves. This, together with the fact that no control group was used, invalidates James's conclusions. Peterson[3] repeated this experiment with nine subjects, using two in the practice and seven in the control. The results indicate clearly the insufficiency of the number of subjects, as well as the faulty method of conducting the practice. In the training series one subject lost about as much

[1]Thorndike, Educational Psychology, 1913, Vol. 2, p. 365.
[2]William James, Principles of Psychology, 1908, Vol. 1, pp. 667-668
[3]*Psychological Review*, 1912, 19, 491-492.

as the other gained, and yet gained more in the tests. The two together, however, gained in the tests more than twice as much as the control group. In the training the two subjects alternated as experimenter and subject, one training with "The Coming of Arthur" and the other with "Guinevere." Thus, while each subject was primarily trained in memorizing one of the poems, he was incidentally trained with the other. These conditions seemed to facilitate the progress of one subject and interfere with the other; the latter subject made greater progress in the tests. Surely the results of such an experiment cannot be accepted as valid.

The best known and most elaborate of the earlier investigations of the transfer effects of memory-training was that of Ebert and Meumann.[1] They sought an answer to the following question: "Is there a general memory function which can be perfected upon any material involving the use of memory, or, on the other hand, must we posit related or unrelated special memories"? Six subjects were tested with a wide range of material involving the use of memory. They were then trained in memorizing four lists of twelve nonsense syllables a day for eight days. After which they were retested with a series of tests similar to those used before the practice. There was another eight days of training with nonsense syllables, followed by a third series of tests. The results indicate considerable improvement in both the training and the tests. The authors conclude that "there are no doubt related memory functions which can be perfected upon any material involving the use of memory, the development taking place proportionately to the degree of relationship between the practice and the test material." In addition to certain irregularities in conducting one of the tests, the validity of every test has been called in question. There is no guarantee that the three test series were of equal difficulty. Too few subjects were used—only six—and in three out of the seven tests in immediate learning there were only two. However, the fact that no control group was used is sufficient to invalidate the conclusions of the authors. Dearborn[2] sought to remedy this defect by repeating the tests only, and concluded as follows: "The results indicate that a considerable part of the improvement

[1]Ueber einige Grundfragen der Psychologie der Uebungsphänomene im Bereiche des Gedächtnisses. *Archiv. fur die gesamte Psychologie*, 1904, 4, 1-232.

[2]W. F. Dearborn, *Psychological Bulletin*, 1906, 6, 44.

found must be attributed to direct practice in the test-series, and not to any 'spread' of improvement from the practice series proper. There is further, at times, lack of correlation between the amount of improvement made in the practice and that made in the test series; occasionally a larger percentage of gain is made in the latter than in the practice itself. This again indicates the presence of direct practice in the test-series. Some at least of the remaining general improvement found is to be explained simply in terms of orientation, attention, and changes in the technique of learning. These results seem to render unnecessary the hypothesis proposed by Ebert and Meumann to account for the large extent of the general influence of special practice, which their experiments seem to indicate."

The most elaborate of the more recent investigations of the transfer effects of memory-training is that made by W. G. Sleight,[1] of England. The general plan was as follows: Three series of ten memory tests were prepared in advance, and as far as possible the three series were equally difficult. The subjects were divided into four groups. One group took the tests only, and was known as the control group, or group number one; the other three were trained in memorizing. Group number two was trained with poetry; group number three with arithmetical tables, and distances from London to the chief towns of England, etc.; and group number four with prose substance. The training occupied thirty minutes a day, four days in a week, over a period of six weeks. The four groups were tested before the practice began, and again at the end of three weeks, and finally at the conclusion of the practice. Sleight says: "Care was taken that each group should work under similar conditions, and that the group which underwent no memory training was never allowed to have the impression that it was in any way handicapped or under conditions not similar to those of the others." The tests were as follows:

1. Points in Circles; an adaptation of a test used by McDougall and Burt.

2. Dates. Two series, each consisting of six dates and their corresponding events, were repeated by the subjects after the experimenter a given number of times. The event was then read out, and the subjects wrote the date.

3. Nonsense Syllables. The syllables were printed in

[1] *British Journal of Psychology*, 1911, 4, 386-457.

white chalk upon a blackboard disc which revolved at a constant rate behind a screen in which there was a rectangular opening through which the exposure was made. Each of the eight couplets was exposed five times, the subjects repeating them aloud as they appeared, with emphasis upon the second syllable. The experimenter repeated the first syllable of each pair and the subjects wrote the associated syllable.

4. Poetry. A stanza of from eight to twelve lines was read to the class; after which the group repeated each line after the experimenter. After several such repetitions (Sleight does not say how many), the subjects wrote what they could remember. This was covered and the stanza given a few more repetitions, followed by another attempt to reproduce it in writing. The correct items in each attempt constituted the test result.

5. Literal Prose. The method of procedure in this test was the same as that in the poetry test; the repetitions being for the first attempt six, and for the second, three.

6. Prose Substance. The selection was read twice to the subjects; after which they were told to write the substance.

7. Map Test. A large map of the world was exposed. Each subject had a corresponding outline map on his desk. The experimenter as he indicated a position on the wall map would announce its name. The wall map was then covered, the name announced, and the subjects required to locate with a cross the position on their own maps. Forty positions were used; the first sixteen being given out two at a time, and the remaining twenty-four, three at a time.

8. Dictation. A prose selection divided into intelligible and grammatically complete portions, beginning with eight and increasing gradually to nineteen words. Each portion was dictated once. The subjects immediately wrote what they remembered.

9. Letters. Consonants only were used. There were 16 series; the first and second contained four letters each; the third and fourth five letters each; the fifth, sixth, seventh, and eighth, six letters each; the ninth, tenth, eleventh, and twelfth, seven letters each; and the remaining four, eight letters each. Each series was dictated once, the subjects immediately reproducing it in writing.

10. Names. Forty-four common Christian names and surnames were used, dictated first in two series of two pairs

each, then in eight series of three pairs each, and lastly in four series of four pairs each. After the experimenter had read two pairs, or three pairs, or four pairs, he repeated a surname and the subjects wrote down the corresponding Christian name. The names were not given in the order in which they were first read. The subject's score was determined by the number of correct names.

In only two of the ten tests were errors scored, and these were arbitrarily penalized. Attention is called to the following precautions used in giving the tests:

1. All answers were written.
2. Plenty of time was allowed for every answer.
3. No test was begun or carried on unless every subject appeared to be giving attention.
4. Every unfamiliar test was preceded by a short practice.

The subjects were 84 Sixth Standard girls of average age 12 years, 8 months.

Sleight says that no attempt was made to estimate numerically the direct effect of the practice, because the conditions under which the practice took place did not admit of this. He simply assumes that improvement occurred. The three kinds of practice were conducted orally. In training with poetry the experimenter would read a line and the subjects would repeat it after him. From 20 to 30 lines were used in this way each day until the average child could repeat them without help. The same procedure was used in practicing with arithmetical tables. In the training with prose substance the selection was read twice to the subjects who then reproduced the substance in writing. In the final results Sleight indicates three scales of certainty of significance: First, those where the superiority or inferiority of a practice group is at least five times the probable error; second, where the difference is at least three times the probable error; and third, where the difference is between two and three times the probable error. The remaining numbers he considers of no significance whatever. According to this method of determining the significance of the results, the practice groups show neither superiority nor inferiority to the control group in four tests: Dates, Poetry, Letters, and Names. In the other tests there are some significant differences; in some cases positive, in other cases negative. The significant cases are indicated in the following table, or Table A.

TABLE A.

Practice Groups	Tests	Superiority over Control	P.E.
Poetry	Nons. Sylls.....	66	11
	Prose Subst.....	–22	11
	Map Test.....	50	14
	Dictation	–32	12
Tables	Points in Circles.....	48	18
	Nons. Sylls.....	85	11
Prose Subst.	Literal Prose.....	21	11
	Prose Subst.....	31	11

An examination of Table A shows that the Poetry Practiced have a superiority in the Test with Nonsense Syllables, and in the Map Test; and an inferiority in the Tests with Prose Substance, and Dictation. The Tables Practiced have a superiority with the Test Points in Circles, and Nonsense-Syllables. The Prose Substance Practiced have a superiority in the Tests with Prose-Substance, and Literal Prose.

A short time after the completion of the above investigation Sleight conducted a somewhat similar experiment with Women Students of the average age 18-19. There were four groups as before. Group number one took the tests only, while groups 2, 3, 4, practiced with Poetry, Tables, and Prose-Substance, respectively. Instead of ten tests there were only six. They were as follows:

1. Dates. A series of ten dates, each with its associated event was repeated six times, after which the subjects reproduced the dates as the events were announced.

2. Nonsense Syllables. The procedure was the same as that used in the first experiment. However, there were twelve couplets instead of eight.

3. Poetry. The subjects repeated after the experimenter line by line a stanza containing about eighty words. The repetitions were preceded by one complete reading of the poem by the experimenter.

4. Prose (Literal). One complete reading of a prose extract by the experimenter was followed by four repetitions on the part of the subjects according to the method used with the poetry.

5. Prose-Substance. A prose extract was read twice by the experimenter. The subjects then reproduced the substance in writing.

6. Letters. This test differed from the Letter Test used in the first experiment in that it was extended to nine letters at one dictation.

The training was carried on for twelve consecutive days except Sunday, for half-an-hour each day. After which the second series of tests was given. The training differed from that of the first experiment. Selections of verse were handed to the subjects of group number two, and they were told to memorize for thirty minutes by whatever method they pleased. A similar plan was followed with group number three. The material used included population, import and export tables, coinage systems, and other similar data of a somewhat irregular form. In the training with prose substance the procedure was the same as that in the first experiment, except in the length and difficulty of the selections. The results which Sleight regards as significant are presented in Table B.

TABLE B.

Practice Groups	Tests	Superiority over Control	P.E.
Poetry	Nons. Sylls.....	33	13
	Poetry	33	16
Tables	Dates	59	24
Prose Subst.	Nons. Sylls.....	–62	13
	Prose Subst.....	52	31
	Letters	–27	13

The significant results in the two experiment are summed up in Table C. This will enable the reader at a glance to compare them.

TABLE C.

	Younger Group			Older Group		
Practice Groups	Tests	Superiority over Control.	P.E.	Tests	Superiority over Control.	P.E.
Poetry	Nons. Sylls....	66	11	Nons. Sylls.....	33	13
	Prose Subst....	–22	11	Poetry	33	16
Tables	Nons. Sylls....	85	11	Dates	59	24
Prose Subst.	Literal Prose..	21	11	Prose Subst....	52	31
	Prose Subst....	31	11	Letters	–27	13
				Nons. Sylls.....	–62	13

Sleight proceeds to interpret these results and the differences in the transfer effects between the two experiments

by using the introspections of the older subjects. The Poetry Practiced in both experiments have a superiority in the Tests with Nonsense Syllables. While the method of training differed, the method of testing was the same. The element of rhythm was emphasized, for in repeating the syllables the accent was on the second syllable of the couplet. With the younger group poetry training interfered with the Prose-Substance Test, and with the older group failed to affect it. This was probably due to the fact that the method of training the younger subjects emphasized rhythm more than the individual methods used by the older group. This is shown in the superiority of the younger group in the Test with Nonsense Syllables, which is 66, with a probable error of 11, while for the older group it is only 33, with a probable error of 13. Any tendency to apply rhythm to the reproduction of the gist of a prose passage would very likely cause interference. With the older group training with poetry facilitated the test with poetry, but failed to affect it with the younger group. This may have been due to the fact that while the method of testing was the same for the two groups, the methods of training were different. Sleight says that the poetry used in the practice was chosen for its simplicity, while that in the tests was not so chosen, some of it being unnecessarily difficult. He mentions other facts which probably "led to a different distribution of attention."

The Tables used with the younger group contained the element of rhythm, while those used with the older group did not. This together with the fact that the method of training the younger group admitted of emphasizing rhythm, while with the older group each subject memorized the tables as she pleased, probably accounts for the fact that the Tables Practiced younger group has a superiority with Nonsense Syllables while the older Tables Practiced group seems to be unaffected in this test. On the other hand, the older group has a superiority in the Dates Test, while the younger group appears unaffected. According to Sleight the difference lies in the fact that the older group instead of making use of rhythm as the younger group did, employed a special kind of visualizing power in the practice with the tables which they were able to use in the Dates Test.

Practice with prose substance had a positive effect upon the Prose-Substance Test with both groups. However, with

three other tests there was a difference. With the younger group Nonsense Syllables were unaffected, while with the older group there was interference. After examining the introspections of the older group Sleight reaches the conclusion that the interference was caused by the strongly contrasted feeling-tone accompanying the exercise, while with the younger group this feeling-tone was absent. He says: "One was auditorially presented, the other visually; one consisted of connected logical speech, the other of disconnected meaningless words; one was arhythmic, the other rhythmic." In spite of the fact that the older group has an inferiority of 27, with a probable error of 13 in the test with Letters, Sleight concludes thus: "With regard to other subjects, the group practicing prose substance reproduction remains in the same position it occupied in the first cross-section." It seems to the writer that the differences in age of the two groups, together with the fact that the methods of practice and the methods of conducting two of the tests were very different, are sufficient to account for the differences in results between the two groups.

Sleight is to be commended for conceiving and executing a memory experiment so elaborate. No one appreciates more than the writer the many difficulties involved. Although the conclusions may be true, yet one would feel surer of them if the practice had been conducted under measurable conditions and identical methods had been used in both investigations. It is always hazardous to leave it to subjects to practice as they please. Some of his conclusions are very interesting. They may be summed up as follows:

1. Specific memory-training is specific in its results.
2. In some cases practice precludes new adjustments.
3. There is no general memory function which can be sharpened upon any material.
4. Differences in the midst of great similarity in the mental processes involved may lead to a loss of "Transfer" or even to reciprocal interference.

G. C. Fracker[1] sought to discover the transfer effects of memory-training in an ingenious and unique experiment by so arranging the test series and the practice series "that the elements concerned in the transference might be determined by analysis of the final results." He had twelve subjects—

[1]*Psychological Review, Monograph* 38, 56-102. 1908.

eight in the practice group and four in the control. The two groups were tested as follows:

1. Memorizing stanzas of poetry.
2. Memorizing the order of four shades of gray according to the method used in the practice series.
3. Memorizing the four intensities of sound used in the practice series, but in a series of nine instead of ten.
4. Memorizing the four grays used in the second test, but in a series of nine.
5. Memorizing four tones (major chord on piano) presented according to the method of the practice series.
6. Reproduction of nine geometrical forms exposed all at once on a card for ten seconds.
7. Reproduction of nine two-place numbers in their order after one hearing lasting 13.5 seconds.
8. Memory of extent of arm movements.

The practice consisted in memorizing four different sound intensities of the same tuning-fork. The four intensities were presented in a certain order, each intensity lasting a half-second, with an interval of a half-second before the next intensity was presented, and so on until the series of four had been presented. Then followed an interval of four seconds, after which the four intensities were presented as before except in a different order. During the next interval of four seconds the subject had to recall the order of the first four intensities presented; then followed another presentation of the four intensities in a still different order; after which the subject recalled, during the interval of four seconds, the order of the presentation in the second series. Thus the training proceeded. One subject took 3600, six took each about 3000, while one took only six hundred of these trials. The two groups were then re-tested with the eight tests used before the practice began.

This investigation has been thoroughly analyzed and discussed by Sleight,[1] in England, and Thorndike[2] in this country. Between the two, the possibilities of criticism have been so exhausted that hardly anything is left to the writer except to indicate the defects pointed out by them. Thorndike, by omitting from consideration two subjects of the practice group—one who did worse at the end of the training, and one who

[1] *British Journ. of Psychology*, 1911, 4, 395-398.

[2] Educational Psychology, 1913, Vol. 2, 393-396.

trained only two days—reaches the conclusion that "on the whole it seems safe to say that a gain in the peculiar ability to grasp the order of four un-named facts by naming them or otherwise, and to hold them while reporting a similar previous set and grasping another, carries over from sound intensities to pitches and grays to half of its own amount and improves the grasping and holding of a series of nine grays, forms or numbers to one-fifth of its own amount—all the tests being subject to the same general conditions of a laboratory experiment." Attention is called also to the fact that the two observers mentioned made great gains in the test series most closely allied to the practice.

Sleight calls attention to several irregularities, and especially to the fact that the withdrawal of subject number five, who made rather spectacular gains in six of the tests, reduces the average percentage of improvement of the trained over the untrained from 19 and 10, to 13 and 10, in the test with the nine grays. He points out also the fact that by omitting the subject of the control group who showed a large retrogression in some of the tests, the practice group has a decided superiority in only two tests—the four grays and the four tones. The facts pointed out plainly indicate the insufficiency of the number of subjects used, and, consequently, the unreliability of some of Fracker's conclusions. On the basis of his subjects' introspections Fracker concludes as follows: "We are able to say that transference depends upon the nature of the imagery employed in the practice rather than upon any other factor." "If in the mind of the observer, the imagery is capable of adjustment to different tasks, it can be used in both improvement and transference, for the elements of the training act are thereby made the same as those of the test act. If it is adapted, in the mind of the observer, to the training task only, it may assist in improvement, but it may interfere with transference."

Mr. W. H. Winch, of England, has made several experiments on the problem of transfer. Most of them deal with memory functions. Since none of them is satisfactory, and so many are similar, a discussion of the more typical and important will suffice. One[1] of the first has to do with the effect of memorizing poetry upon rote memory for historical passages and descriptions about places. Although the practice was very

[1] *British Journ. of Psychol.*, 1908, 2, 284-293.

brief—from forty to sixty minutes plus the time for reproduction—there was considerable improvement. In the final tests the practice group gained much more than the control group. However, it should be borne in mind that it was rote memorizing in both the practice and the tests.

In another experiment[2] Winch investigated the transfer effects of training in rote memory upon substance memory. The practice group was trained 20 minutes a day for three days—one day in each successive week—in rote memory. The control group was occupied during the same period in drawing difficult geometrical designs. The two groups were then retested with prose substance. The practice group gained 21 per cent. and the control group 10 per cent. in the tests, while the practice group gained 13 per cent. in the training. Winch arbitrarily assumes that 2 per cent. of this was due to growth, and concludes as follows: "About as much or more improvement reckoned in percentages, as has been made in the practice medium itself—rote memory for meaningless things—has been transferred to substance memory." The results are astonishing when one considers the very limited amount of training the practice group received. In the tests less than six lines of very simple prose was read three times to the subjects. This would constitute rote memory for many of the subjects. Then, too, scoring such tests accurately is always difficult. At least two or more disinterested judges ought to score such material independently. Such an experiment, instead of establishing anything conclusive about transfer, rather accentuates the importance of greater care in conceiving and conducting a transfer experiment. The very fact that the superiority of the practice group is about the same as its improvement in the training series, together with the fact that very simple test material was read three times to the subjects, leads one to suspect that the function tested was the function trained. With another group Winch made a somewhat similar experiment. Instead of one practice period in each of three weeks, there was one in each of thirteen weeks. The superiority of the practice group was only 6 per cent. Taking the two experiments together, one is inclined to say that the more practice there is in rote memory the less improvement there is in substance memory.

Winch also investigated the transfer effects of practice in

[2] *Ibid.*, 1909-1910, 3, 386-405.

substance memory upon productive imagination with school boys of average age 13 years, four months.[1] The practice consisted in giving the substance of short stories which the subjects had studied visually for five minutes. The tests in productive imagination consisted in producing a story which included certain words given to the subjects. The average improvement in the prose substance practice was 37 per cent., while the superiority of the practice group in producing stories from the words given was 15 per cent. Winch concludes: "We know then that an improvement in imagination has resulted from an improvement in memory due to practice; but we do not know what percentage of the improvement in memory due to practice has been transferred." In a second experiment of this kind with school girls of average age 12 years and eleven months the practice group trained until it ceased to improve and gave evidence of being tired of the work. The results indicate an inferiority of 12 per cent. on the part of the practice group, which leads Winch to conclude as follows: "The view that memory can be overtrained and thus have prejudicial results on imaginative work is confirmed." He holds that there is a community of function between substance memory and productive imagination, and yet insists that a limited amount of training in the practice has a positive effect, while overtraining has a negative effect. Scoring in such experiments is not only very difficult but often accompanied by a constant error. He says also that the teacher who administered the tests and exercises in the second experiment "was a strong believer in the cultivation of memory functions, and was not aware that I intended to endeavor to produce a decline in the inventive function." The very fact that Winch endeavored to bring about certain results in the second experiment may have influenced him in the selection of the material, in the procedure, and in the scoring. If three competent judges had scored the tests in both experiments without knowing to which group a subject belonged, the results might have been different. It seems rather strange that a small amount of practice should have produced a superiority of 15 per cent., and much practice an inferiority of 12 per cent. It will be recalled that in another experiment three practice periods produced a superiority of 13 per cent., and in a similar experiment 13 practice periods caused a superiority of 6 per cent. Surely no one is justified

[1] *British Journ. of Psychol.*, 1911, 4, 95-125.

in venturing a conclusion on the basis of these experiments, except to point out the need of greater caution.

The most comprehensive and influential of the earlier investigations of transfer was made by Thorndike and Woodworth.[1] They examined:

1. The influence of certain special training in the estimation of magnitudes on the ability to estimate magnitudes of the same general sort.
2. The influence of training in observing words containing combinations of letters or some other characteristics, on the general ability to observe words.
3. The influence of special training in memorizing on the general ability to memorize.

In a part of these investigations the authors made use of for the first time a control group. Their results and conclusions were so radical that widespread interest in the problem was aroused, and, as a consequence, many investigations were inaugurated. Possibly the most significant and striking part of their conclusions is to be found in these words: "Improvement in any single mental function rarely brings about equal improvement in any other function, no matter how similar, for the working of every mental function-group is conditioned by the nature of the data in each particular case." In his recent survey of the literature on the problem of transfer, Professor Thorndike has this to say: "The change is simply the necessary result upon the second function of the alteration of those of its factors which were elements of the first function. Training is not totally general; neither is it totally specialized."[2]

One of the most interesting experiments in connection with this problem is the one on the perception of illusions by Judd.[3] The method was that of the Müller-Lyer Illusion. There were two observers, one of whom was Judd, who says he "was trained somewhat irregularly and with a background of abstract knowledge and expectation." He practiced with a figure 54mm. long, oblique lines at 90° with each other, keeping the standard on the right, both figures being held in a horizontal position. The illusion which was 44.8 mm. (i. e., this was the judged length) at the first trial disappeared after 980 trials, "not by any process of judgment or of indirect correction, but

[1] *Psychological Review*, 1901, 8, 247-261, 384-395, 553-564.
[2] Educational Psychology, 1913, vol. 2, 359.
[3] *Psychol. Review*, 1902, 9, 27-39.

by coming to look differently than it did at first." In the test before practice the illusion of figure 54mm., 45°, was 45.1mm., after practice, 53.0mm. With figure 68mm., 90°, before practice 58.7mm., after practice, 67.3mm. In these tests it should be remembered that the figures were kept in a horizontal position with the standard on the right. Again using figure 54mm., 90°, 175 determinations were made with the standard on the left. At first the illusion was almost as strong as in the beginning of the practice series. However, the curve rose rapidly, the illusion not quite disappearing with the 175th trial, "thus indicating clearly," says Judd, "the transfer of practice." Turning the figures into a vertical position with the standard above, and again with the standard below, gave almost perfect results in 80 trials. A comparison was made also between the standard 54mm., 90°, and a simple straight line placed somewhat below and to the left. Judd insists "that in spite of the varying conditions there was a transfer of practice." Observer E. was especially prepared for a study of the inversion of the standard figure. He was tested with figure 53mm., 90°, the standard being on the left. The illusion was 45.6mm. E. then practiced with figure 68mm., 45°, with the standard on the right. After 750 trials he was again tested with figure 53mm., 90°, the standard at this time being on the right. The illusion was 50.4mm. After the whole practice series of 950 trials was completed, the illusion having not quite disappeared, he was tested again with figure 53mm., 90°, the standard being on the left, as in the original test. Professor Judd at this point emphasizes the fact that "E. did not know anything about the disappearance of the illusion in the practice series, and that no information was given him in regard to the change in figure or the difference in position of the standard line. The results were astonishing. The illusion was 41mm., and not even 1500 trials made any marked change in it." Judd says: "In spite of change in the length of the lines perceived, and in spite of a change in the degree of obliquity of the additional lines, and, finally, in spite of a new arrangement of the figures, the effects of the practice were obviously carried over to the new conditions. Furthermore, it is evident from the curve that the observer started on this second series of measurements with a thoroughly established habit of interpretation. The subject who had cultivated in a purely empirical way this fixed mode of interpretation, misapplied it,

to be sure, and by this misapplication of the mode of interpretation be increased the strength of the illusion under the new conditions."

Scholkow and Judd[1] sought to discover what effect an understanding of the principle of refraction would have upon practice in hitting a target placed under water at different depths. One group of boys in grades 5 and 6 practiced with such knowledge, while another practiced without it. When the target was twelve inches under water, both groups did equally well. However, when the depth was changed to four inches the informed group readily adjusted themselves, while the uninformed group was much confused, their errors being persistent. It is evident that practice in hitting a target at a depth of 12 inches caused facilitation in the one case and interference in the other when the depth was changed.

Kline[2] tested seventeen subjects in marking parts of speech in English prose. Eight of these subjects were retained as a control group, while nine were trained in cancelling e's and t's, from 30 to 45 minutes daily, for fourteen days. The time-limit method was used, the scores being in number marked per minute. The group made considerable improvement, the gains ranging from 31 per cent. to 168 per cent. in the number marked per minute. When the two groups were retested, both made gains, but the control group gained the more. In other words the practice caused interference in the tests. Of all the stimuli presented in the practice situation, only e's and t's were to be cancelled, while such stimuli as parts of speech were to be neglected. In the test situation the conditions were reversed; the stimuli to be neglected in the practice situation were now to be cancelled, while the stimuli determining the cancelling in the practice situation were now to be neglected. Kline offers the following explanation: "The meaning of the relatively inferior work of the practice group is best made out from the reports of the members of the group. One says: (1) 'In crossing out parts of speech one always had to think what part of speech the word was.' (2) 'The crossing out of the letters became a habit and instead of crossing out words one wanted to cross out e's and t's. These seemed to be seen so much more clearly than the parts of speech.' Another writes, 'The practice with e's and t's hindered me in dealing with the parts of

[1] *Educational Review*, 1908, 36, 28-42.

[2] Bulletin of the State Normal School, Duluth, Minn., Feb., 1909.

speech. I think it was because I was accustomed to looking for e's and t's and the tendency was to cross out those letters rather than the parts of speech.' "

Coover and Angell[1] tested four subjects in discriminating shades of gray, each test consisting of thirty-five judgments, made on three separate days. Three other subjects took the tests also, but made their judgments on two separate days. The four subjects were trained in discriminating sound-intensities for seventeen days, over a period of fifty-seven days, making forty judgments a day. In the tests the judgments were given in the categories of "lighter," "darker," "like," and "undecided." In the training the judgments were made in terms of "louder," "softer," "like," and "doubtful." At the conclusion of the training the control group of three and the practice group of four were re-tested in discriminating shades of gray. Results are given in per cents. The practice group gained 32 per cent. in right judgments, while the control group lost 7 per cent. About all that one can make out of these results is that the practice group did a little better in judging differences of brightness after training in judging differences of sound-intensities, while the control group did worse in the second seventy judgments in discriminating brightness than in the first seventy. How reliable the figures are is not stated. The authors conclude as follows: "That efficiency of sensible discrimination acquired by training with sound stimuli has been transferred to the efficiency of discriminating brightness stimuli, and that the factors in this transfer are due in great part to habituation and to a more economic adaptation of attention, i.e., are general rather than special in character." The investigation is not reported in sufficient detail to enable one to go into a thorough analysis of the procedure so as to test the reliability of the results. The very fact that the control group did worse in the second test should be regarded as significant, for it plainly indicates too few subjects and, perhaps, some irregularity in giving the tests.

Coover and Angell[2] report also an experiment in which they sought to discover the effect of practice in card-sorting upon typewriter reactions. There were four subjects in the practice group and three in the control. The training consisted of about fifteen exercises in sorting cards, distributed over a

[1] *American Journ. of Psychol.*, 1907, 18, 327-336.
[2] *Ibid*, 336-340.

period of forty days. Previous to the training the practice group was tested for five days in typewriter reactions, and after the training, for three days; while the control group was tested for only three days before and two days after. To make the results comparable both groups should have been tested in exactly the same way and under the same conditions. Since there is no common basis of comparison conclusions based on such an experiment are unreliable. The practice group improved more in the tests from the first three days before training to the last three days before training, than from the last three days before training to the three after training. Investigations of this kind, when conducted with so few subjects and so little regard for scientific accuracy, only emphasize the necessity and importance of consistent, accurate procedure. They certainly do not justify any such conclusion as that "training the activity of Reaction with Discrimination and Choice by sorting cards into compartments has increased the facility of a like activity in both speed and regularity in 'typewriter-reaction' (a) noticeably, in two cases, after the latter had become automatic, and (b) markedly in two others, in the course of practice."

Gilbert and Fracker[1] sought to determine what effect practice in reaction and discrimination with stimuli in one sense has on the same process in other senses, the latter not having been practiced at all. There were three subjects, each of whom was tested in reacting simply, and reacting "with choice," to the following stimuli: Color, pressure, and electricity. Each day thereafter for twelve days two of the subjects practiced reacting simply, and also "with choice" to sound. The other subject practiced eleven days in reacting to simple sound only. The three were then tested again as in the beginning. In the opinion of the authors the results justified the following conclusions:

a. That practice in reaction to sound reduces the time of reaction to other forms of stimuli by amounts almost equal to the reduction of the time of sound reaction itself.

b. That such practice alone does not reduce the time of discrimination and choice.

c. That practice in discrimination of sounds reduces also the time of discrimination for other forms of stimuli.

The second conclusion is based on the fact that the subject

[1] *University of Iowa Studies in Psychology*, 1897, 1, 62-76.

that practiced in simple reaction only, made scarcely any gain in reaction "with choice." The figures show that his record in the first test was almost as good as the records of the other two in the second test. Consequently, we should expect him to make little improvement, for the first test shows him to be much nearer the physiological limit than either of the others. There were too few subjects and not sufficient practice to make tne conclusions of any scientific value. Then, too, the functions tested and the functions trained contained many identical elements.

Bair[1] made several experiments on the influence of practice in forming certain associative habits upon the ability in certain different habits. We shall notice two that bear somewhat upon the problem of transfer.

1. Taking six symbols, letters or figures, Bair made a series of fifty-five; with six different symbols he made another series of fifty-five; and so on until he had twenty sets of these series. Having labeled six keys of a typewriter with the six symbols of the first series, he exposed the fifty-five symbols of this series, in chance order, one by one, while the subjects upon seeing a symbol tapped the corresponding key. The time required to tap out the series was recorded. In the same way six other symbols were used with the series composed of them. The time required to tap out this series was recorded. This was kept up until the twenty sets had been used. It will be noted that the symbols were changed from test to test, thus changing somewhat the conditions of the experiment. Four subjects took the tests. The first improved from 62 to 52; the second, from 95 to 85; the third, from 71.5 to 58; and the fourth, from 65 to 56. Bair claims that the major part of this gain could not have been due to merely getting used to the machine or the general features of the experiments, for the fourth subject was already used to these and still gained nearly as much as any one of the others. It is hardly necessary to call attention to the fact that the conditions of each succeeding test were the same, with the exception of the six symbols. Thus there were many identical elements, and these were being practiced from test to test. Probably the slight improvement made was due to the practice of these elements.

2. The other experiment "consisted in taking daily records, for twenty days, by means of a stop-watch, of the time

[1] *Psychol. Review*, Monogr. 19, 1902, pp. 25, 28, 64-67.

required to repeat the alphabet from memory. Each day's experiment was as follows: First, the alphabet was repeated as rapidly as possible forward; secondly, the letter n was intercepted between each of the letters; thirdly, the alphabet was repeated as rapidly as possible backward; and lastly, the alphabet was repeated backward intercepting n between each of the letters. At the end of twenty practices in each order, the subject repeated the alphabet, first, forward, intercepting, instead of n, the letter x, and repeating three times; secondly, intercepting r, and repeating three times; then lastly, repeating backward, and in like manner intercepting x and r and repeating three times." In the test series there was an improvement equivalent to that of three days' practice in the training series. Here, as before, there were many identical elements.

Ruger[1] in an "experimental study of the processes involved in the solution of mechanical puzzles and in the acquisition of skill in their manipulation," also studied the transfer effects, and attempted to isolate the factors involved. The materials and method of procedure are too intricate for an intelligible and detailed description without some study of the puzzles themselves. With the aid of his subjects' introspections in connection with objective measurements Ruger made a classification of the transfer factors into (1) general factors, and (2) special factors. He says: "It has seemed advisable to the writer to use the term transfer in a very broad sense to include the effect of any given experience on any subsequent one whether the effect results directly or by means of an idea, whether the transfer is one of method, or of material, or of motor processes, and whether it is positive or negative." And, again, "to the writer the problem consciousness considered in itself and as to conditions of efficiency seems to have many characteristics in common irrespective of the degree of relatedness of the material concerned."

Foster[2] studied the effects of practice upon visualizing and upon the reproduction of visual impressions, and reached the conclusion: "That training in these experiments has made the observers noticeably better observers or memorizers in general, or given them any habits of observing closely or reporting correctly, or furnished any ability to meet better any situations generally met with, neither we nor the observers themselves believe."

[1] *Archives of Psychology*, 1910. No. 15.
[2] *Journal of Educational Psychology*, 1911, 2, 11-21.

Whipple[1] tested the effect of practice upon the range of visual attention and of visual apprehension and concluded that the improvement was due to "habituation to experimental conditions" and the use of "assimilative devices."

In his "Educative Process," page 42, Bagley reports an experiment, made by Squire, on the transfer effects of special training in neatness in arithmetical work, on neatness in other school work. In his report Bagley says: "At the Montana State Normal College careful experiments were undertaken to determine whether the habit of producing neat papers in arithmetic will function with reference to neat written work in other studies; the tests were confined to the intermediate grades. The results are almost startling in their failure to show the slightest improvement in language and spelling papers, although the improvement in the arithmetic papers was noticeable from the very first."

On the other hand, Ruediger[2] reports that neatness cultivated in connection with one school subject did improve neatness in other subjects. However, it would seem almost impossible for a teacher to follow the instructions given without suggesting to the pupils the importance of neatness in other school subjects as well as the activities specially mentioned. Attention is called to two of the instructions in which the writer thinks these suggestions are involved.

"1. Talk freely **with** the class (not to) on the importance of neatness in dress, business, the home, hospitals, etc., connecting it as far as you can with the subject under experiment. Guard against overdoses.

"2. Do **not** bring up the subject of neatness with the other studies of the school. If the pupils bring up these studies quietly substitute something else. Talk of neatness only in that class, not to the school in general."

The writer feels confident that the fact that neatness was demanded in one school subject, and emphasized as an ideal in everything else than the other school subjects, aroused suspicion on the part of the pupils, and perhaps, was more effective in influencing their behavior in the tabooed subjects than if such subjects had not been so conspicuously isolated and neglected.

[1] *Journal of Educational Psychology*, 1910, 1, 249-262.
[2] *Educational Review*, 1908, 36, 364-371.

Jastrow[1] found Hermann and Keller, sleight-of-hand performers, not quite as good in reaction-time experiments as the average college student. Quick as they were in performing their tricks, they were not quick enough to escape detection when required to perform them in slightly different ways.

Raif[2] tested expert pianists and found that they could not manipulate a telegraph key any more rapidly than persons of average intelligence.

Thorndike[3] calls attention to the fact that experiments on the amount and rate of improvement with practice "shows negatively that the practice in the tasks of school and life which an earnest graduate student has had leave him still far below his possibilities—so far below that a very small amount of time devoted to any special function improves it greatly."

In presenting this survey the writer has followed the same general plan used by Thorndike in his excellent presentation and discussion of the investigations of the problems of transfer in his Educational Psychology, Vol. 2, chapter 12. Although the writer acknowledges his indebtedness for the free use of Thorndike's presentation, yet, wherever possible, the presentations and discussions are the writer's own. In concluding his discussion Thorndike says: "These experimental facts as a whole leave a rather confused impression on one's mind, and resist organization into any simple statement of how far the improvement wrought by special practice spreads beyond the function primarily exercised. They do, however, at least put out of court the old doctrine of a very wide spread of a very large percentage of the special improvement." To the present writer these investigations emphasize also the necessity and importance of a standard method of procedure to which investigations of the problem of transfer should conform. Surely sufficient experimentation has been done to make clear what constitutes a standard method of procedure. Thorndike very aptly says: "The experimental facts now at hand are in a sense trivial in comparison with the very great variety of facts which must be measured in order to describe justly the spread of improvement in the work of the schools, trades, and the like."[4]

[1] *Science, N. S.*, 1896, 3, 685-689.
[2] *Zeitschrift fur Psychologie*, 1900, 24, 352-356.
[3] *Educational Psychology*, 1913. Vol. 2, 415-416.
[4] *Op. cit.* Vol. 2, p. 417.

CHAPTER II.

DESCRIPTION AND DISCUSSION OF MATERIALS AND PROCEDURE

Section 1. Preliminary Investigations.

In three preliminary investigations the writer sought to discover, within a limited field, an answer to some of the questions involved in the problem of the transfer effects of practice.

A. Reaction Time

The first experiment had to do with reaction time. Three subjects were tested in reacting to a soft sound rather than a loud one; to a light pressure rather than a heavy one; and to blue rather than red. They were then practiced for eight weeks in reacting each 200 times daily to a sound of medium intensity. At intervals of two weeks the three tests were repeated. There was considerable improvement in the test series, but not so much as in the practice series. It was impossible to determine how much of the improvement in the test series, if any, was due to transfer, as there was no control group by which to measure the improvement due to direct practice in the tests themselves.

B. Cancellation

The second experiment was made with cancellation tests. Nineteen subjects were engaged—nine for the control group, and ten for the practice group. Both groups were tested at the same time, by the time limit method, with five cancellation tests. The practice group was trained 30 minutes a day for 10 days in cancelling in English prose words containing both a and t. On account of eye strain four of the subjects were forced to withdraw from the practice. When the practice was completed, the two groups—six of the practice group and nine of the control group—were reassembled and retested with the five tests used before the practice began. In only one test did

the practice group give reliable evidence of superiority to the control group, and that was the test in Spanish prose, in which words containing both a and t were cancelled. The elements which determined the cancelling in this test were the same as the elements which determined the cancelling in the practice series.

C. Cancellation

In the third experiment the materials of the practice series and of the test series were the same as those used in the second experiment. In fact these two experiments were in progress at the same time. The subjects in the second were boys and girls in the fifth grade of the Speyer School, New York City; while the subjects in the third were girls in the Scudder Secretarial School of New York City. They ranged in age from 18 to 19 years. Thirty-five girls were tested with the five tests. Fourteen of them volunteered to practice regularly every day as they had time. At the end of three weeks only seven had practiced, while only three of these had done the amount required. However, these seven were accepted as the practice group. The two groups were reassembled and tested as before. The practice group had a decided superiority over the control group in the Spanish a—t Word test. In the other tests there were no reliable differences as measured by the probable error.

Section II. The Present Investigation.

As a result of these preliminary experiments, seven cancellation tests, to be described later, were decided upon; and in order that the investigation might be thorough the writer determined to secure as large a group of subjects as possible under conditions which would enable him to carry the experiment through to a successful conclusion. The general plan has been set forth in the Introduction. Briefly, it was as follows: Having decided that the practice should consist in cancelling, in English prose, words contained both a and t, seven cancellation tests more or less similar to the practice material, were selected. Eighty boys, ranging in age from 11 to 13 years, of the Hebrew Orphans Home,[1] of New York City, were tested

[1]This Institution is more like a large boarding school than like an orphanage. Individuality in manner and dress, spontaneity and freedom of expression, are manifest characteristics of the children. The Superintendent and one of the Governors cooperated with the writer in arousing interest in the experiment and in conducting the practice and tests with systematic regularity.

twice with each of these tests. Thirty-six of these eighty subjects were trained for four ten-minute periods a day during sixteen days in cancelling a—t words in English prose. At the conclusion of the practice all of the subjects (with the exception of four) were retested twice with each of the seven tests used before the practice.

A. Practice Material

The material for the practice was selected with special reference to the problem to be solved. One hundred copies of a pamphlet called the "Gold Standard," published by Ginn and Co., in 1896, were used in the practice series. These books contained 143 pages of prose printed on excellent paper. There were no pictures and the reading matter was of no special interest to the subjects.

B. Test Material

The test materials were selected with special reference to the task involved in the practice.

1. The Spanish a—t Word Test. This test consisted of a paragraph of Spanish prose in which all words containing both a and t were to be cancelled.

2. The Spanish e—s Word Test. This test consisted of the same paragraph of Spanish prose used in Test 1, but now the words containing both e and s were to be cancelled.

3. The a and t Letter Test. This test consisted of a distribution of small letters in which the letters a and t were to be cancelled.

4. The e and s Letter Test. This test consisted of the same distribution of small letters used in Test 3, but now the letters e and s were to be cancelled.

5. The Columbia A Test. This test consisted in a distribution of capital letters in which the A's were to be cancelled.

6. The Thorndike B Test. This test consisted in a distribution of capital letters different from that used in Test 5, in which the B's were to be cancelled.

7. The Woodworth and Wells Number-Group Checking Test. This test consisted in a distribution of six-place number-groups, in which groups containing both 4 and 7 were to be cancelled.

C. Procedure

Each subject was supplied with a drawing pencil, which had been properly sharpened. Before the blanks were distributed one was shown to the subjects and the task to be performed was carefully explained. The subjects were then asked questions concerning the task, and, in turn, were invited to ask questions of the writer, if they did not understand clearly what the task was. They were told that the test blanks would be distributed face down and that they were to write their names on the back of them, and that just three seconds before they were to begin cancelling, the signal "ready" would be given, which would mean that the blanks were to be turned over and pencils held in readiness for the command to "start" cancelling. They were also told that the cancelling would last exactly one minute, and that at the signal "stop," they were to lay the pencils on the table. Promptly at the conclusion of each test the blanks were collected and another set distributed. The same method was used with each one of the seven tests. Immediately upon the completion of the series, the seven tests were given again in the same order with the same instructions as before. The subjects were urged at the beginning of each test to cancel only the designated symbols and to cancel as rapidly as possible.

All of the work, both the practice and tests, was done at night between 7:15 and 8:30, in one of the large school rooms of the Hebrew Orphans Home. The seven tests described above were given on October 22, 1914 and will be known in this investigation as the Initial Test Series. On October 26 thirty-seven of the eighty subjects began the practice, and will be referred to as the Practice Group; while the remaining subjects will be known as the Control Group.[1] The practice periods were each 10 minutes long and there were four of them each night for 16 nights, distributed over a period of 22 days. There was no practice on Friday and Sunday night. At the beginning of each practice period, the task to be performed was carefully explained and the subjects urged to do their best in speed and accuracy. Each subject was provided with a drawing pencil properly sharpened. These pencils along with the

[1]As a matter of fact, on the 12th day of the practice, one subject of the practice group was expelled on account of cheating; and three of the control group failed to take the final test, on account of sickness. Consequently the practice group consisted of 36 and the control group of 40 subjects.

books were collected promptly at the conclusion of the practice each night.

Precautions were taken against cheating. One of the Governors of the Home, and two trustworthy monitors, remained in the room during the practice. Promptly upon the signal from the writer to stop cancelling, the monitors, under his direction, rubber-stamped the place reached by each subject. The books were closed and the pencils placed on the desk. Careful watch was kept to see that no subject tampered with his book during the relaxation interval allowed between the periods. On the 12th night of the practice one subject opened his book during this interval and began cancelling. He was promptly expelled with the approval of those in authority. The interval between practice periods was from six to eight minutes.

Realizing the importance of securing a maximum of improvement in the practice, the following incentives were employed: (1) On the second night of the practice, two prizes were offered, one to the subject making the greatest improvement, and one to the subject cancelling the largest number of words with the fewest errors. (2) Two entertainments were promised; one during the practice, and another at the conclusion if the practice proved satisfactory. (3) Before beginning the practice each night, individual scores of the previous night were read to the group, and each subject urged both to excel the other subjects and to beat his own record. (4) During each relaxation interval, the subjects were given candy or entertained with short stories or music. (5) During each practice period, the writer was constantly among the subjects encouraging them by his presence and interest. In the meantime, the members of the control group were assured repeatedly that they were just as capable and important as the practice group. They were promised an entertainment if they did their best in the final test. Their behavior indicated an interest in the experiment equal to that of the practice group. It should be said that these promises were fulfilled. Both groups were fully informed as to the nature of the investigation.

The practice was concluded on Monday night, November 16. On the following night, the two groups—the control group of forty and the practice group of thiry-six—were reassembled, and again tested with the same seven tests used before the practice began, each test being used twice. All of the material,

both of the practice series and of the initial and final test series, was scored carefully by the writer himself. The scoring consisted in counting the number of cancellations and errors, the errors being mostly those of omission. The total number of a—t words cancelled in the practice series was 452,995 and the total number of errors was 35,719. The data of the experiment will be found in the next section.

Since more than 99 per cent. of the errors were errors of omission, the total ground covered, or gross speed, is represented by the sum of the cancellations and errors. Consequently the per cent. of accuracy per period can be obtained by dividing the number of cancellations per period by the total ground covered, or gross speed. For instance, the average number of cancellations for the first practice period was 102.6 a—t words, and the average number of errors (mostly omissions), 27 a—t words. Therefore, the total ground covered, or gross speed, was 129.6 a—t words. By dividing 102.6 by 129.6 the per cent. of accuracy is obtained.

Section III. Treatment of the Data

A. Data of the Practice Series

Instead of presenting in detail all of the data of the practice series, the data of the group performance in each practice period will be presented. It will be found in Table I.

The reader should bear in mind that each practice period was ten minutes long, and that there were four of them each night for sixteen nights. Consequently, there are sixteen first-periods, sixteen second-periods, sixteen third-periods and sixteen fourth-periods. The table is so constructed that one can not only follow the progress of the group both in speed and accuracy of cancellation through each successive practice period, but from each first-period to the next first-period, and from each second-period to the next second-period, and so on with the sixteen third-periods and sixteen fourth-periods. The first double column contains the data for the sixteen first-periods; the first half containing the average number of cancellations, and the second half the corresponding percents of accuracy. The remaining columns of the table are constructed in exactly the same manner.

If one wishes to follow the progress of the practice group in successive practice periods, he reads across the table thus: In

the first period the group has an average performance of 102.6 cancellations, and an accuracy of 79 per cent.; in the second period, 101.0 cancellations and 82 per cent. of accuracy, and so on to the fourth period. In the fifth period (the first period of the second day) an average of 132.3 cancellations and an accuracy of 86 per cent., and so on through the table. If, on the other hand, one cares to follow the progress in each suc-

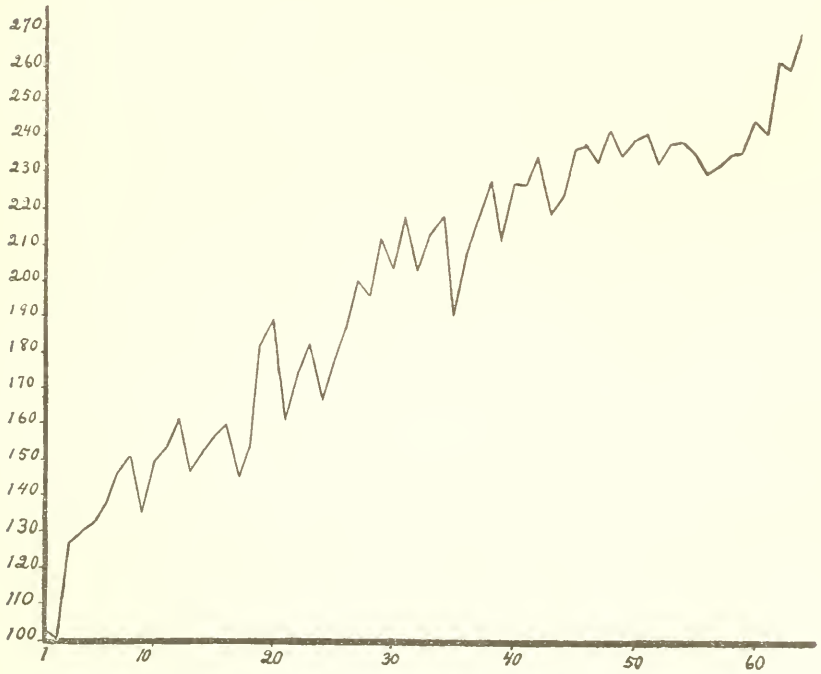

Figure I—Showing Progress of the Practice Group in Average Number of Cancellations per Period, through the Sixty-four Practice Periods.

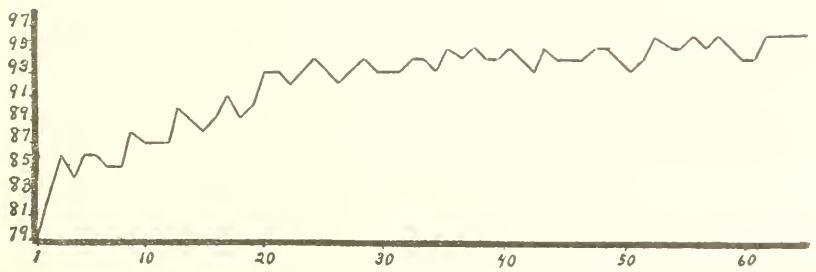

Figure II—Showing Improvement in the Percent of Accuracy.

TABLE I.

DATA OF THE PRACTICE SERIES

(The figures represent the average for the thirty-six subjects. Each of the four periods per day is averaged separately and occupies a separate column in the table. Accompanying each average is the corresponding per cent. of accuracy for that period.)

Days	First Period		Second Period		Third Period		Fourth Period	
	Av. Can.	%Acc.	Av. Can.	%Acc.	Av. Can.	%Acc.	Av. Can.	%Acc.
1	102.6	79	101.0	82	126.4	86	128.9	84
2	132.3	86	137.7	86	145.8	85	151.3	85
3	134.9	88	149.1	87	153.4	87	161.2	87
4	146.0	90	152.2	89	156.6	88	159.2	89
5	145.1	91	153.0	89	182.2	90	189.1	93
6	160.9	93	174.3	92	182.0	93	166.3	94
7	176.2	93	186.1	92	196.1	93	192.6	94
8	211.8	93	203.2	93	218.4	93	202.8	94
9	212.9	94	216.8	93	190.0	95	206.1	94
10	216.6	95	226.4	94	211.1	94	225.5	95
11	225.2	94	234.3	93	218.3	95	223.4	94
12	235.2	94	236.2	94	232.3	95	241.3	95
13	234.6	94	238.3	93	240.1	94	231.9	96
14	236.5	95	237.0	95	234.6	96	228.9	95
15	231.3	96	233.6	95	234.2	94	243.4	94
16	239.5	96	260.7	96	257.6	96	266.5	96

ceeding first-period, or second-period, and so on, he reads down the double columns. If the practice is followed from period to period as it occurs and the changes observed, one will notice that out of the sixty-three possible changes there are forty-four in which there is an increase in the number of cancellations; ten of these are accompanied by an increase in accuracy, twenty by a decrease and fourteen by no change. There are eighteen cases in which there is a decrease in the number of cancellations, two of which are accompanied by a decrease in accuracy, thirteen by an increase and three by no change. The one case in which there is no change in the number of cancellations is accompanied by a decrease in accuracy. As the writer deemed it unwise to attempt to get introspections from subjects so young, he has no explanations of these facts to offer from the point of view of the subjects themselves. However, he noticed after the first week of practice, that the subjects were striving more and more for accuracy. During the second and third weeks more than half of the errors were made by fewer than eight subjects.

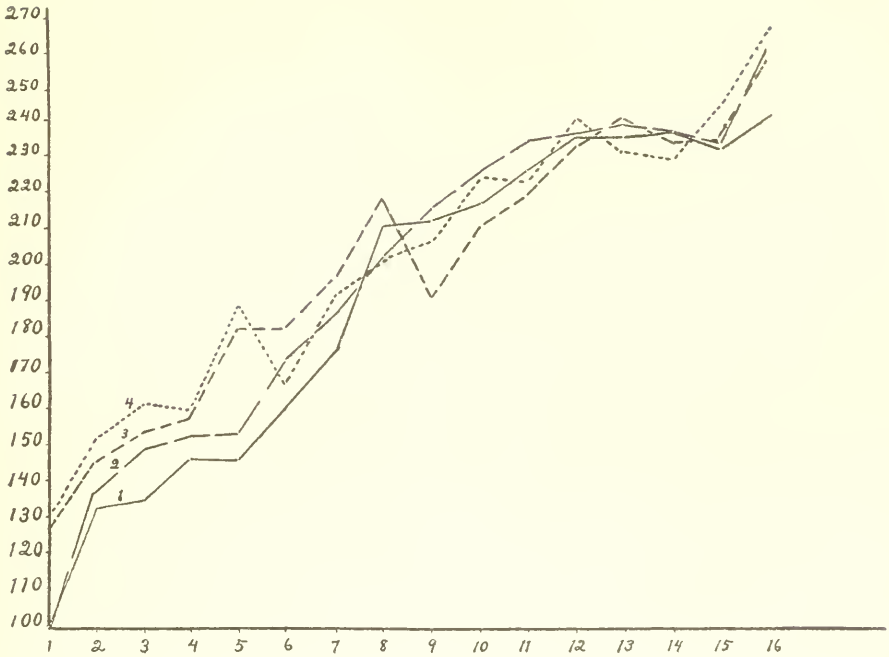

Figure III—Showing Improvement in Speed of Cancellation in Terms of the Four Daily Practice Periods.

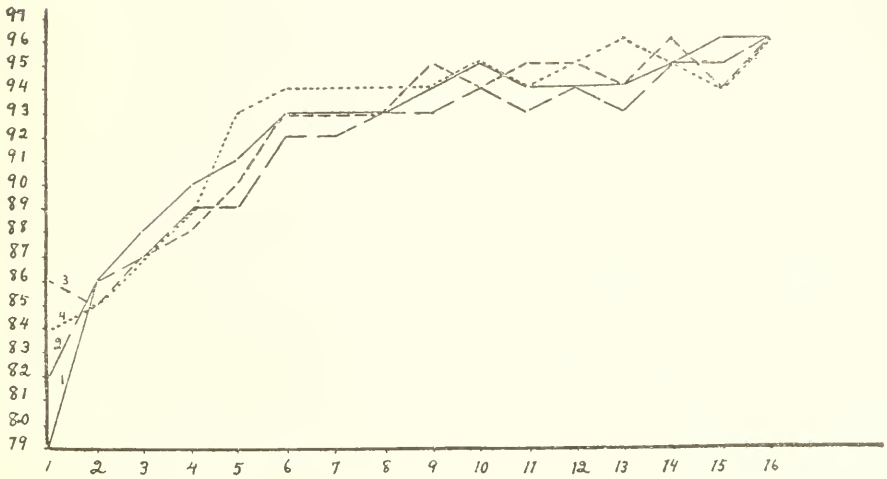

Figure IV—Showing Improvement in Percent. of Accuracy in Terms of the Four Daily Practice Periods.

The data in Table I are presented in graphic form in figures I, II, III and IV. Figure I shows the progress of the practice group in average number of cancellations per period for the sixty-four ten-minute periods of the practice series, while figure II shows the corresponding per cents. of accuracy. The numbers along the horizontal axis indicate the practice periods, while those along the vertical axis indicate, in figure I, average number of cancellations per period, and in figure II, per cents. of accuracy. In figure III the progress in average cancellation per period is presented, not through each successive period, but through each of the sixteen first-periods, each of the sixteen second-periods, and so on. The continuous line represents the progress through the first-periods, the broken line through the second-periods, the dash line through the third-periods and the dotted line through the fourth-periods. The corresponding per cents. of accuracy are presented in the same way in figure IV. An inspection of these figures shows that the practice group made considerable progress both in speed and accuracy of cancellation. This improvement is evident both in the graphs showing the consecutive practice periods, and also in those showing the four separate practice periods of each day in separate curves.

B. Data of the Test Series

Table II presents in detail the data of the seven tests for the practice group. In the first column the subjects are indicated by numbers 1 to 36. Under Spanish a—t Word Test are four columns; the first two contain the data of the initial test series, while the last two contain the data of the final test series. Under "Can." will be found the number of actual cancellations per minute for each subject, and under "Er." the corresponding errors. The data of the other six tests are recorded in exactly the same way. It should be borne in mind that in each series of tests—the initial and the final—each of the seven tests was given twice, and that the numbers recorded in Table II are the averages of these two tests. The table reads thus: In the Spanish a—t Word Test subject No. 1 has an initial performance of 7.0 cancellations per minute and 1.5 errors, and a final performance of 15.0 cancellations per minute and 2.0 errors; in the Spanish e—s Word Test an initial performance of 10.0 cancellations per minute and 1.0 errors, and a final performance of 11.0 cancellations per min-

ute and 0.0 errors, and so on with each of the other tests. The performances of the other subjects are read in the same way.

TABLE II.

DATA OF THE PRACTICE GROUP IN THE TEST SERIES

	Spanish a-t Word Test				Spanish e-s Word Test			
	Initial		Final		Initial		Final	
Sub.	Can.	Er.	Can.	Er.	Can.	Er.	Can.	Er.
1	7.0	1.5	15.0	2.0	10.0	1.0	11.0	0.0
2	9.0	1.5	14.0	0.0	9.5	2.5	11.5	1.5
3	8.0	0.0	16.0	0.5	10.5	0.0	12.0	3.0
4	9.0	1.5	17.5	3.5	11.0	0.0	12.0	4.0
5	9.5	1.0	19.5	0.0	11.0	1.0	11.5	3.0
6	12.5	1.0	19.5	1.5	11.5	0.0	15.0	1.5
7	10.0	3.0	15.5	1.0	10.0	1.5	8.5	3.0
8	11.5	0.0	18.0	0.0	12.5	0.0	12.5	1.0
9	11.0	0.0	18.0	4.0	11.5	0.0	14.0	9.0
10	11.5	0.0	23.0	1.5	11.0	0.5	16.0	7.0
11	11.0	1.0	17.5	1.5	11.0	1.0	10.0	5.0
12	11.5	0.0	19.5	0.5	12.0	0.0	12.0	1.5
13	10.5	1.5	19.5	1.0	12.0	0.5	12.5	4.0
14	10.5	1.0	18.5	2.0	11.0	0.0	10.0	3.5
15	9.0	1.0	18.0	1.0	10.5	0.0	11.0	2.0
16	9.0	2.5	19.5	0.5	10.5	0.5	11.5	1.0
17	11.5	1.0	25.5	1.0	15.5	2.5	20.5	8.0
18	10.0	0.0	18.5	0.0	10.5	0.5	12.0	1.0
19	8.5	2.0	20.0	0.5	10.0	0.5	12.5	3.0
20	10.0	0.5	19.5	1.5	11.0	0.0	15.5	4.5
21	9.0	2.0	14.0	2.5	10.5	1.0	10.0	10.5
22	11.0	0.5	17.0	3.0	12.0	2.5	12.0	4.5
23	11.5	0.5	19.0	2.5	12.0	0.0	17.0	3.5
24	9.5	1.5	20.0	1.0	7.5	4.0	15.0	7.5
25	10.5	0.5	17.5	3.5	11.0	0.5	12.5	7.0
26	9.5	6.5	20.5	2.5	11.5	1.5	11.5	2.0
27	8.0	2.5	15.5	0.5	11.5	1.5	12.0	5.0
28	7.0	4.5	17.0	1.5	11.0	2.0	10.5	4.5
29	10.5	1.0	21.5	1.5	11.5	0.5	16.5	3.5
30	11.0	1.5	20.5	0.5	14.5	6.0	15.0	10.5
31	11.0	0.5	17.5	0.5	11.0	1.0	12.0	2.5
32	10.0	1.5	16.5	0.5	10.0	1.0	9.5	3.0
33	9.5	0.5	14.0	0.5	10.0	0.0	11.0	0.0
34	12.0	12.0	21.0	4.5	12.5	1.0	11.0	10.0
35	9.0	0.5	20.0	3.5	10.5	1.0	11.5	16.0
36	11.5	0.0	21.5	1.5	11.5	0.0	13.0	2.5

This table shows the initial and final performance of each member of the practice group in each of the seven tests, in terms of number of cancellations and errors per minute.

TABLE II (Continued).

DATA OF THE PRACTICE GROUP IN THE TEST SERIES

	Letter Test a and t				Letter Test e and s			
	Initial		Final		Initial		Final	
Sub.	Can.	Er.	Can.	Er.	Can.	Er.	Can.	Er.
1	27.0	0.0	34.0	1.0	40.0	0.0	42.0	1.5
2	29.0	5.5	29.5	2.5	42.0	1.0	41.5	1.5
3	18.5	2.5	29.5	2.0	36.0	0.0	39.5	2.5
4	26.5	8.0	30.5	0.0	39.0	4.0	49.0	1.0
5	32.0	6.0	40.0	6.5	47.5	3.0	53.5	0.0
6	31.5	2.0	40.0	1.0	44.0	1.0	52.5	2.0
7	27.5	6.0	38.5	4.5	40.0	2.5	41.5	1.0
8	27.5	3.5	37.5	0.0	45.5	0.5	41.0	2.0
9	27.0	8.5	38.0	1.0	43.0	0.0	52.5	1.0
10	30.5	6.0	48.0	0.5	47.5	2.5	56.0	4.0
11	25.0	1.5	30.5	1.0	37.5	0.5	39.0	0.5
12	30.5	2.5	47.5	1.0	40.5	0.0	49.5	3.0
13	27.5	9.0	36.5	1.5	41.0	4.5	42.0	0.5
14	24.0	0.5	33.0	1.0	43.5	1.5	32.0	3.0
15	18.5	3.0	26.0	1.5	30.5	1.5	29.5	0.0
16	32.5	2.0	48.0	1.0	53.5	2.0	57.0	1.5
17	29.0	1.0	46.0	0.5	53.5	1.5	53.0	2.5
18	26.5	3.0	42.0	1.0	39.0	2.5	47.0	0.5
19	24.5	1.0	35.0	3.0	36.5	1.5	41.5	2.5
20	24.0	1.5	42.5	2.5	37.0	1.0	48.0	2.0
21	21.0	10.5	32.5	6.5	29.5	8.0	33.5	10.0
22	24.0	2.5	23.5	1.5	37.5	0.0	29.0	1.5
23	34.0	4.5	41.0	4.5	47.5	1.0	46.0	0.5
24	33.5	5.5	34.5	2.5	44.0	4.0	45.0	2.5
25	20.5	0.0	36.0	1.0	34.5	3.0	39.5	2.0
26	36.0	4.5	39.5	2.5	39.0	3.0	34.0	1.5
27	25.0	6.0	41.5	5.0	44.0	1.5	42.0	6.5
28	25.5	4.0	30.5	4.0	39.0	1.0	45.0	3.0
29	30.0	5.5	42.0	1.5	45.5	1.5	52.5	2.5
30	43.5	5.5	53.0	1.5	70.5	2.0	66.5	2.5
31	31.5	1.0	35.5	0.0	36.5	1.0	36.0	2.0
32	24.0	0.5	26.0	3.5	32.0	0.5	26.0	4.0
33	18.0	4.5	29.5	3.0	28.0	2.5	30.0	0.5
34	36.0	11.0	50.5	1.5	46.5	7.5	55.0	7.5
35	21.0	1.5	36.5	6.5	36.0	5.0	40.0	10.5
36	26.0	3.0	40.0	3.0	38.5	1.5	39.0	2.0

TABLE II (Continued).

DATA OF THE PRACTICE GROUP IN THE TEST SERIES

	Capital A Test				Capital B Test			
	Initial		Final		Initial		Final	
Sub.	Can.	Er.	Can.	Er.	Can.	Er.	Can.	Er.
1	36.5	1.5	45.0	0.5	18.5	0.5	20.0	0.5
2	37.0	1.0	40.0	0.0	20.5	3.0	21.0	1.5
3	34.0	3.0	42.5	0.5	15.5	0.0	22.0	2·5
4	40.5	0.5	46.5	0.0	19.0	0.0	18.0	1.0
5	47.0	1.0	62·5	0.5	24.0	0.5	31.5	1.0
6	41.5	0.0	61.0	0.0	23.0	0.5	31.0	3.0
7	41.0	0.0	54.5	1.0	22.0	0.0	31.0	0·5
8	40.0	0.0	56.0	1.0	20.0	1.0	25.0	4.0
9	36.5	0.0	46·5	1.0	19.5	1.0	25.5	0.0
10	32.0	3.5	72.0	5.0	20.5	0.5	32.0	4.0
11	37.0	3.0	43.5	0.0	20.0	1.0	25.0	0·0
12	41.5	3.0	57.0	0.0	23.0	0.5	34.0	0.0
13	41.5	0.0	51·0	0.5	21.5	0.5	32.0	0.5
14	38.5	0.0	48.5	0.0	19.5	0.0	20.0	0.5
15	33.5	1.0	48.0	0.0	14.5	0.0	19.0	2·5
16	44.5	1.5	62.5	0.5	24.5	0.5	31.5	0.0
17	37.5	0.0	61·5	1.0	21.5	0.0	28.0	2.0
18	37.0	0.0	49.0	0.0	20.0	0.5	28.5	1.0
19	41.5	0.5	57.5	0.5	16.5	0.5	22.0	7·5
20	30.5	1.0	47.0	4.5	17.5	1.0	29.5	3.5
21	37.0	2.0	42·5	3.0	17.0	3.5	19.0	1.5
22	35.5	0.5	37.0	0.5	25.5	0.0	24.0	0.0
23	38.0	1.0	48.0	1.5	20.0	0.0	22.0	7·0
24	43.5	0.5	49.0	0.0	20.5	0.5	28.5	4.5
25	35.5	0.0	44·5	0.0	16.0	0.0	21.0	0.5
26	49.5	0.0	58.5	0.0	25.5	1.5	22.0	5.0
27	31.5	1.5	50.5	6.5	19.5	1.5	25.0	5.0
28	35.5	0.0	50·5	2.0	18.0	2.0	23.0	1.5
29	38.0	1.0	57.5	0.5	23.0	0.5	31.0	1.5
30	45.5	8.0	60.5	0.5	33.5	0.5	34.0	1·0
31	40.0	0.5	52.5	0.0	23.5	0.5	27.5	1.0
32	35.0	0.0	41·5	0.5	19.5	1.0	24.0	0.0
33	32.5	1.0	44.0	0.0	18.5	1.0	22.5	0.5
34	41.0	6.0	66.5	3.5	32.5	4.0	33.0	3·0
35	34.5	0.0	49.0	0.5	21.5	0.0	28.5	1.0
36	38.5	0.5	59·0	0.5	18.5	0.5	24.5	2.0

TABLE II (Continued).

DATA OF THE PRACTICE GROUP IN THE TEST SERIES

	4 and 7 Group Test			
	Initial		Final	
Sub.	Can.	Er.	Can.	Er.
1	11.0	0.0	16.0	1.0
2	18.0	0.5	16.5	1.0
3	9.0	0.0	17·5	0.0
4	14.5	0.0	18.0	2.5
5	16.5	0.0	21.0	0·5
6	19.0	0.0	22.5	0.5
7	14.0	0.0	16.5	2·0
8	11.0	0.0	18.0	0.5
9	16.0	0.5	20.0	2.0
10	20.0	0.5	29.0	0.0
11	13.0	0.0	17·5	5.0
12	19.0	0.0	25.0	0.0
13	18.0	0.5	17.0	1.0
14	16.0	0.0	15.0	7.0
15	9.0	0.0	17.5	0.0
16	21.0	0.0	28.0	0.0
17	21.5	1.0	34.0	1.0
18	11.0	2.5	16.0	1.0
19	15.5	1.0	24·0	1.0
20	14.5	1.0	22.0	5.0
21	10.0	0.5	15.5	4.0
22	20.0	0.0	19.5	1.5
23	15.5	0.0	22.0	1·0
24	12.0	0.0	19.0	2.0
25	11.0	0.5	16.5	4.5
26	13.0	0.0	20.0	0.0
27	14.5	0.5	17·5	3.0
28	15.5	1.0	19.5	2.0
29	14.0	1.0	19.0	0.5
30	29.0	0.5	34.5	0.5
31	18.0	0.5	20.0	0·0
32	14.5	0.0	20.0	0.0
33	9.5	0.0	14.5	0.0
34	26.0	0.5	26.0	4.5
35	16.0	0.0	24.0	3.0
36	12.5	1.0	18.5	1.0

Table III gives in detail the data of the Control group in the two test series. This table is constructed in exactly the same way as Table II. In the first column the subjects are indicated by the numbers 1 to 40. It reads thus: In the Spanish a—t Word Test Subject No. 1 has an itial performance of 10.5 cancellations per minute and 0.5 errors, and a final performance of 12.0 cancellations per minute and 0.0 errors; in

TABLE III.

DATA OF THE CONTROL GROUP IN THE TEST SERIES

	Spanish a–t Word Test				Spanish e–s Word Test			
	Initial		Final		Initial		Final	
Sub.	Can.	Er.	Can.	Er.	Can.	Er.	Can.	Er.
1	10.5	0.5	12.0	0.0	11.0	0.0	11.5	0.0
2	10.5	1·0	11.0	0.0	7.5	4.5	11.0	0.0
3	10.0	0.5	13.5	0.0	10.5	0.0	12.5	0.0
4	6.5	3.0	11.0	2.5	11.0	1.5	14.5	0.5
5	10·0	2.0	11.0	0.0	10.5	0.5	11.5	0.0
6	10.5	1.0	13.0	1.0	10.0	1.5	13.0	1.0
7	8.5	1·5	10.0	1.0	11.5	0.5	11.5	0.0
8	9.5	1.5	10.0	0.0	11.0	0.0	10.0	0.0
9	10.5	0.5	12.0	0.5	10.5	0.0	11.0	0.0
10	11·0	1.5	11.5	1.5	12.0	0.0	12.0	1.5
11	8.0	1.0	10.5	0.0	11.0	0.0	11.0	1.0
12	9.5	0.5	11.0	0.5	9.5	0.5	10.5	1.0
13	10.5	1·5	12.0	0.5	12.0	0.5	12.5	0.0
14	10.0	1.0	12.0	1.0	12.5	0.5	14.5	0.0
15	7.5	2.5	11.5	1.5	10.0	2.0	11.5	2.5
16	9·5	0.5	12.0	0.0	10.0	0.5	11.5	0.0
17	7.0	1.0	9.0	0.5	9.0	2.5	10.5	0.0
18	10.0	0.0	11.0	1.5	10.5	0.5	10.0	2.0
19	12.0	1·0	12.5	0.5	12.0	1.5	12.0	0.0
20	10.0	0.0	10.5	1.5	8.0	0.0	11.0	0.0
21	8.5	1.5	10.5	0.5	10.0	1.5	10.5	0.5
22	10·0	0.0	11.5	0.0	11.0	0.5	12.0	0.0
23	7.5	1.5	13.0	2.0	11.0	0.5	15.0	4.5
24	10.0	0.0	11.5	2.0	12.5	0.0	15.0	0.5
25	9.5	0·5	12.5	0.0	10.0	0.5	11.0	0.0
26	10.0	0.0	11.0	0.0	10.5	1.0	10.5	1.0
27	9.0	1.0	8.5	6.0	10.0	1.0	10.5	2.0
28	9·5	2.0	10.0	0.5	11.0	0.0	12.0	0.0
29	7.5	0.5	11.0	0.0	8.5	0.5	10.5	0.0
30	9.5	1.0	11.0	1.0	10.0	1.5	12.0	0.0
31	12.0	0·5	13.0	0.5	11.5	0.0	12.5	1.0
32	10.0	0.5	11.5	1.5	11.0	0.0	12.0	0.0
33	7.5	1.5	10.5	0.5	9.5	1.5	10.0	1.0
34	2·5	7.0	9.0	0.5	10.0	1.0	8.5	4.5
35	6.5	4.5	9.5	0.0	6.5	4.5	12.0	11.0
36	8.0	1.5	10.0	0.0	11.0	0.0	11.0	0.0
37	10.5	0·5	12.0	0.0	12.0	0.0	15.0	0.5
38	10.0	0.5	14.0	2.0	14.0	1.5	18.0	1.5
39	9.5	1.5	9.5	2.0	11.0	0.0	11.5	0.0
40	10·5	0.5	13.0	1.0	14.0	0.0	16.5	0.5

This table shows the initial and final performance of each member of the control group in each of the seven tests, in terms of number of cancellations and errors per minute.

TABLE III (Continued).

DATA OF THE CONTROL GROUP IN THE TEST SERIES

	Letter Test a and t				Letter Test e and s			
	Initial		Final		Initial		Final	
Sub.	Can.	Er.	Can.	Er.	Can.	Er.	Can.	Er.
1	29.5	1.5	32.5	1.5	41.5	1.5	38.0	4.0
2	28.5	1.0	26.5	2.5	36.5	0.0	38.0	0.5
3	22·5	1.0	34.5	1.5	36.5	0.5	42.0	0.0
4	29.0	4.5	34.0	2.0	46.5	1.5	54.5	1.0
5	19.5	0.5	22.0	0.0	32.0	0.5	35.0	0.0
6	29.0	1·0	37.5	2.5	40.0	0.5	49.5	1.5
7	32.5	2.5	34.0	3.0	42.0	2.0	55.5	1.0
8	25.0	1.0	24.0	0.5	41.0	1.0	36.0	0.0
9	26·5	0.5	30.5	0.5	38.0	0.5	34.0	0.0
10	32.0	1.5	32.5	0.5	40.5	2.0	41.5	0.0
11	24.5	2.0	32.0	2.5	36.0	1.5	45.0	0.5
12	27.0	0·0	27.0	1.5	38.0	0.5	42.0	0.0
13	29.0	0.0	30.5	2.0	40.5	0.5	49.5	0.5
14	32.5	1.0	38.5	0.5	45.5	2.0	58.0	0.0
15	23·0	5.0	31.0	9.0	37.0	1.0	47.5	0.5
16	25.5	2.0	27.5	0.5	32.5	1.0	33.5	1.5
17	18.5	2.0	16.5	1.0	30.0	1.5	31.5	0.5
18	17.5	2·5	25.0	3.5	32.0	0.0	42.5	0.0
19	41.0	0.5	41.5	1.0	62.5	3.5	63.5	0.5
20	20.0	0.5	23.5	0.0	31.0	1.5	34.0	0.0
21	27·0	1.5	28.0	2.5	36.0	1.5	35.5	1.0
22	22.0	1.5	27.0	2.5	35.5	0.5	41.5	1.0
23	30.0	2.5	39.0	3.0	40.0	2.0	55.5	2.0
24	29.5	1·0	40.5	3.0	55.0	3.0	64.5	1.5
25	25.5	1.0	31.5	3.5	43.5	2.5	42.0	0.5
26	23.0	0.5	26.5	0.0	39.0	0.5	35.0	0.0
27	24·0	3.0	33.5	5.5	35.5	1.0	46.5	0.5
28	20.0	3.5	29.0	7.0	35.0	0.5	45.5	1.0
29	22.5	0.0	31.5	3.5	34.0	1.5	48.0	0.5
30	30.5	6·0	36.0	1.0	41.0	0.5	49.0	2.0
31	25.0	1.0	35.5	0.5	34.0	0.0	39.0	0.0
32	25.0	3.5	29.0	0.0	34.5	2.5	43.0	1.5
33	23·0	6.5	26.0	4.0	33.5	3.0	38.5	1.5
34	19.5	3.5	30.5	2.5	35.0	2.0	36.0	3.0
35	22.0	5.0	31.5	1.5	29.0	4.5	40.0	2.0
36	22.5	0·5	23.5	1.0	34.0	0.5	42.0	0.5
37	24.0	2.5	30.5	0.5	42.0	1.0	45.5	0.0
38	34.5	2.0	44.5	1.5	56.0	3.0	74.0	1.5
39	27·5	4.5	31.5	3.0	45.5	0.0	52.0	3.5
40	31.5	5.5	33.5	2.5	43.5	1.0	61.5	1.0

the Spanish e—s Word Test an initial performance of 11.0 cancellations per minute and 0.0. errors, and a final performance of 11.5 cancellations per minute and 0.0 errors, and so

TABLE III (Continued).

DATA OF THE CONTROL GROUP IN THE TEST SERIES

	Capital A Test				Capital B Test			
	Initial		Final		Initial		Final	
Sub.	Can.	Er.	Can.	Er.	Can.	Er.	Can.	Er.
1	27.5	2.5	42.5	0.0	21.0	0.5	24.5	0.0
2	37.5	0·0	37.5	0.5	20.0	0.5	19.0	0.0
3	31.0	0.5	43.0	0.5	19.5	0.0	28.0	1.0
4	35.5	0.0	48.5	0.5	23.0	2.5	27.0	4.0
5	35·0	0.0	40.0	0.0	19.0	0.0	20.0	0.0
6	40.5	1.0	48.5	0.5	21.0	1.0	28.5	1.5
7	41.5	2.0	40.5	1.5	23.5	1.5	28.5	4.0
8	34.0	1·0	43.5	0.0	19.0	0.5	20.5	0.5
9	44.0	0.0	46.0	0.0	21.5	0.0	20.0	0.0
10	38.5	3.0	39.5	0.5	18.0	0.0	17.0	3.5
11	34·5	0.0	41.0	0.5	16.0	1.0	19.5	0.0
12	36.5	0.0	44.5	0.5	20.5	0.5	23.5	0.0
13	43.5	0.0	47.0	0.5	18.5	0.5	26.0	0.5
14	36.5	0·0	53.0	0.0	20.0	0.0	24.0	0.0
15	31.5	0.5	44.0	0.0	18.5	1.5	21.0	0.5
16	29.5	1.0	36.0	0.0	17.5	1.5	19.0	0.0
17	30·0	0.0	33.5	0.5	15.5	0.5	18.5	0.0
18	39.0	2.0	41.0	0.5	16.5	0.0	22.0	0.5
19	42.5	2.0	57.5	1.0	24.5	0.5	23.0	1.0
20	33·5	0.5	45.0	0.0	17.5	0.0	19.0	1.0
21	29.0	0.0	38.0	0.0	16.0	1.0	17.0	1.0
22	32.0	0.0	42.0	0.0	18.5	0.0	20.0	0·0
23	50.0	0.5	56.0	1.5	25.0	2.5	32.5	0.5
24	39·5	0.5	62.0	1.0	23.0	0.0	33.0	0.0
25	36.0	5.5	42.0	0.0	20.0	0.0	24.0	1.5
26	35.0	0.0	43.5	0.0	19.5	0.5	22.5	0·0
27	34.5	0.5	47.5	0.5	17.0	1.0	19.5	1.5
28	27·0	0.0	38.5	0.5	14.5	4.0	23.5	1.0
29	33.5	0.0	45.0	0.0	15.5	0.5	19.5	1.5
30	46.0	1.5	54.0	1.0	22.5	1.5	26.5	1.5
31	38.0	0.5	47.0	0.0	16.0	1.0	23.5	1.0
32	34·0	1.5	42.5	0.5	18.5	0.5	21.5	0.0
33	32.5	4.5	47.5	0.0	18.0	2.0	20.5	1.0
34	34.5	5.0	46.0	0.5	19.0	0.5	25.5	0·5
35	31.5	1.5	45.0	0.5	17.5	1.5	26.0	5.0
36	30·0	4.5	36.0	1.5	20.5	0.0	21.5	0.5
37	37.5	0.5	46.0	0.0	24.0	0.0	28.0	0.0
38	44.0	3.0	62.0	1.5	27.5	2.0	38.0	1·5
39	39.5	2.5	54.0	1.5	23.5	2.0	29.5	2.5
40	39·0	0.0	55.5	1.0	26.0	0.5	29.5	2.0

on with the other tests. The data for each of the remaining subjects are to be read in the same manner.

TABLE III (Continued).

DATA OF THE CONTROL GROUP IN THE TEST SERIES

	4 and 7 Group Test			
	Initial		Final	
Sub.	Can.	Er.	Can.	Er.
1	17.5	0.0	18·5	0.0
2	16.0	0.0	20.5	0.0
3	13.0	0.0	19.0	0.5
4	12.0	2.0	17.0	1.0
5	14.5	0·0	16.5	0.0
6	18.0	0.0	20.5	0.5
7	16.0	0.5	17.0	0.0
8	11.5	1.0	12.0	0.0
9	17.5	0.0	19·0	0.0
10	17.5	0.5	19.0	0.0
11	15.0	1.0	20.0	0.0
12	15.5	0.0	18.0	0.0
13	16.0	0·5	19.5	0.0
14	18.0	0.0	21.5	1.0
15	16.0	0.0	18.0	0.0
16	11.0	0.0	16.0	0.0
17	9.5	0.0	13·0	0.0
18	14.0	0.5	17.0	0.5
19	22.5	0.0	20.5	0.5
20	11.0	0.0	15.5	0.0
21	14.5	0·0	18.5	0.0
22	14.0	1.0	17.0	0.0
23	17.0	0.5	21.5	0.5
24	16.0	1.0	24.0	0.0
25	18.5	0.0	20·5	1.5
26	14.0	0.0	17.0	0.0
27	12.5	1.0	17.5	1.0
28	9.5	1.0	15.0	1.0
29	10.0	0·0	17.0	0.5
30	15.0	0.0	20.0	0.0
31	17.5	0.5	21·0	0.0
32	15.5	0.5	17.0	0.0
33	13.0	0.5	16·5	0.5
34	8.5	6.0	17.5	0.5
35	11.5	0.5	18.0	2.5
36	14.5	0.5	19.5	0.0
37	18.0	0·0	23.5	0.5
38	24.0	0.0	31.5	1.0
39	14.0	0.5	21.5	1.0
40	15.5	1.0	18.5	0.5

An inspection of Tables II and III shows the differences in performance between the subjects of the control group and those of the practice group. While performance is recorded in cancellations and errors, yet a discussion of change in performance requires that errors be treated with respect to their accompanying cancellation scores. When this is done, we can say whether a subject has changed not only in speed of cancellation but in per cent. of accuracy. Attention is directed to the fact that the change in the control group in both speed and accuracy measures the progress due to direct practice in the tests themselves. Therefore, whatever differences we may note between the performance of the practice group and that of the control group must be due to the transfer effects of the training in the practice series. While a complete discussion and explanation of the transfer effects is reserved for the next chapter, yet it may prove of interest to the reader to consider briefly some of the differences between the two groups without attempting to draw final conclusions. Such a study will indicate what one may reasonably expect when the data are more elaborately treated.

In the Spanish a—t Word Test 80 per cent. of the control group and 72 per cent. of the practice group gained in both speed and accuracy. Fifteen per cent. of the control and 28 per cent. of the practice lost in accuracy while gaining in speed. The gains of the control group range from 5.5 cancellations to a loss of 0.5, and those of the practice group from 5.0 to 14.0. In the Spanish e—s Word Test, on the other hand, 40 per cent. of the control and 8 per cent. of the practice gained in both speed and accuracy, while 5 per cent. of the control and 16 per cent. of the practice lost in both. Twenty-two per cent. of the control and 61 per cent. of the practice lost in accuracy while gaining in speed. Five per cent. of the control and 11 per cent. of the practice lost in accuracy without changing in speed. It is significant that only 32 per cent. of the control and 88 per cent. of the practice lost in accuracy, while 62 per cent. of the control and 69 per cent. of the practice gained in speed.

In the a and t Letter Test 60 per cent. of the control and 80 per cent. of the practice gained in both speed and accuracy, while 2.5 per cent. of the control and none of the practice lost in both. Thirty per cent. of the control and 16 per cent. of the practice lost in accuracy while gaining in speed. Five per cent. of the control and 3 per cent of the practice gained in accuracy

TABLE IV.

SHOWING THE AVERAGES OF THE PRACTICE GROUP IN EACH OF THE SEVEN TESTS BEFORE AND AFTER PRACTICE, TOGETHER WITH THE DIFFERENCES BETWEEN THEM

(This table is derived from Table II)

TESTS	INITIAL TEST SERIES				FINAL TEST SERIES				DIFFERENCES			
	Av. Can.	P. E.	Av.%Acc.	P. E.	Av. Can.	P. E.	Av.%Acc.	P. E.	Av. Can.	P. E.	Av.%Acc.	P. E.
Spanish a–t Words	10.03	0.15	86	1.23	18.47	0.27	92	0.63	8.44	0.31	6	1.38
Spanish e–s Words	11.14	0.12	92	0.88	12.54	0.25	76	1 46	1.40	0.28	–16	1.70
Letters a and t	27.45	0.58	87	0 88	37.35	0.80	94	0 53	9.90	0.99	7	1.02
Letters e and s	41.26	0.77	95	0.47	43.51	1.03	94	0.47	2.25	1.29	– 1	0.66
Capital A	38.33	0.48	97	0.37	51.75	0.95	98	0.28	13.42	1.07	1	0.46
Capital B	20.93	0.40	96	0.48	26.00	0.59	93	0.73	5.07	0.71	– 3	0.87
Number Groups 4 and 7	15.51	0.47	97	0.37	20.47	0.51	92	0.90	4.96	0.70	– 5	0.97

while losing in speed. In the e and s Letter Test 67.5 per cent. of the control and 30.5 per cent. of the practice gained in both speed and accuracy, while 2.5 per cent of the control and 25 per cent. of the practice lost in both. Seventeen and five-tenths per cent. of the control and 33.3 per cent. of the practice lost in accuracy, while gaining in speed. Twelve and five-tenths per cent. of the control and 11 per cent. of the practice lost in speed while gaining in accuracy. The reader will note that 85 per cent. of the control and 63.8 per cent. of the practice gained in speed while 20 per cent of the control and 58.3 per cent of the practice lost in accuracy.

In the Columbia A Test 72.5 per cent. of the control and 67 per cent. of the practice gained in both speed and accuracy. Twenty-two and five-tenths per cent. of the control and 33 per cent. of the practice lost in accuracy while gaining in speed. In the Thorndike B Test 57.5 per cent. of the control and 33 per cent. of the practice gained in both speed and accuracy, while 5 per cent. of each group lost in both. Thirty-two and five-tenths per cent. of the control and 61 per cent. of the practice lost in accuracy, while gaining in speed. Attention is called to the fact that only 37.5 per cent. of the control, and 66 per cent. of the practice lost in accuracy.

In the Number Group Checking Test 75 per cent. of the control and 42 per cent. of the practice gained in both speed and accuracy, while none of the control and 6 per cent. of the practice lost in both. Twenty-five per cent. of the control and 50 per cent. of the practice lost in accuracy while gaining in speed.

Table IV is derived from Table II and shows the average performance in speed and accuracy of the practice group in the initial and final test series. The seven tests are listed in the first column. Under "Initial Test Series" are four columns; the numbers in the first indicate the average number of cancellations per minute, while the numbers in the second represent the corresponding probable errors; in the third and fourth columns the corresponding average per cents. of accuracy with their probable errors are presented. The four columns under "Final Test Series" are similarly arranged. Under "Differences" there are four columns. The first contains the differences between the final averages and the initial averages, while the corresponding probable errors are in the second column. The third column contains the differences between the final and initial per cents. of accuracy, with the corresponding probable errors in the fourth column.

TABLE V.

SHOWING THE AVERAGES OF THE CONTROL GROUP IN EACH OF THE SEVEN TESTS BEFORE AND AFTER PRACTICE, TOGETHER WITH THE DIFFERENCES BETWEEN THEM

(This table is derived from Table III)

TESTS	INITIAL TEST SERIES				FINAL TEST SERIES				DIFFERENCES			
	Av. Can.	P. E.	Av.%Acc.	P. E.	Av. Can.	P. E.	Av.%Acc.	P. E.	Av. Can.	P. E.	Av.%Acc.	P. E.
Spanish a–t Words	9.24	0.17	88	1.09	11.24	0.14	93	0.78	2.00	0.22	5	1.34
Spanish e–s Words	10.62	0.14	93	0.84	12.00	0.18	93	0.95	1.38	0.22	0	1.26
Letters a and t	26.00	0.52	92	0.64	30.99	0.57	93	0.51	4.99	0.77	1	0 81
Letters e and s	39.00	0.68	96	0.27	45.14	1.02	98	0.22	6.14	1.23	2	0·35
Capital A	36.12	0.55	97	0.41	45·55	0.69	99	0.11	9.43	0.88	2	0.42
Capital B	19.81	0.33	96	0·41	23.75	0.59	96	0.45	3.94	0.60	0	0.61
Number Groups 4 and 7	14.87	0.33	96	0.52	18.75	0.29	98	0 26	3.88	0.44	2	0.58

The table reads thus: In the Spanish a—t Word Test the practice group has an initial average performance of 10.03 cancellations per minute with a probable error of 0.15, and an initial average accuracy of 86 per cent. with a probable error of 1.23. In the final test series it has an average performance of 18.47 cancellations per minute with a probable error of 0.27, and an average accuracy of 92 per cent., with a probable error of 0.63. The gain of the practice group is 8.44 cancellations per minute, with a probable error of 0.31, and in accuracy 6 per cent., with a probable error of 1.38. The data for the other tests are read in the same manner.

Table V is derived from Table III and shows the average performance in speed and accuracy of the control group in the initial and final test series. It is constructed in exactly the same manner as Table IV.

CHAPTER III.

INTERPRETATION OF RESULTS

The transfer effects of the practice are to be sought in the superiority or inferiority of the practice group to the control group in each of the seven tests, for it is a valid assumption that the practice group made equally as much improvement because of the direct practice in the tests themselves as the control group. Consequently, whatever difference there may be between the performances of the two groups must be due to transfer effects. In this investigation there are three measures of performance; one measures the speed of cancellation, another the accuracy of cancellation and the third, the gross speed (ground covered or total number of cancellations and errors.) Accuracy of cancellation and speed of cancellation are different and should be treated separately. Some have sought to eliminate errors by arbitrarily penalizing for them; others have simply ignored them. To do either is to ignore the facts in the case. The writer is fully convinced that such procedure would be unscientific in this investigation, for the very truth he is seeking here would be obscured thereby. An exact evaluation of performance requires measures of the characteristics by which performance is manifested, and further, that these measures be treated separately. In all of the tests and in all of the practice, involved in this investigation, the materials were carefully scored for both speed of cancellation and accuracy of cancellation. Consequently, in discussing the transfer effects of practice as revealed in the superiority or inferiority of the practice group the writer is prepared to indicate whether the superiority or inferiority is one of speed of cancellation or of accuracy of cancellation, or of gross speed, or any two or more of these characteristics of performance.

While the writer was primarily engaged in investigating the transfer effects of the practice and not the practice *per se*, yet the practice was conducted with the utmost caution and the highest regard for scientific results. The thirty-six

subjects practiced diligently for four 10-minute periods a day for 16 days under conditions and with incentives highly favorable to improvement. It is well to keep in mind in what the practice consisted. Only words containing a and t were to be cancelled. All other words and letters were to be neglected. Therefore, strengthening the habit of cancelling a—t words meant strengthening the habit of neglecting to cancel all other words. The stimuli to reaction were a—t words irregularly distributed amongst numerous other words; the reaction consisted in recognition of the stimuli and the cancellation of the same. The improvement as measured in terms of speed and accuracy of performance indicates progress in both recognition and cancellation.

By reference to Table I in the preceding chapter the reader may observe the changes of the practice group in measures of speed and accuracy in the practice series. From the first period to the sixty-fourth period the group advanced from an average performance of 102.6 cancellations per period and an accuracy of 79 per cent., to an average performance of 266.5 cancellations per period and an accuracy of 96 per cent. No matter how the improvement is estimated, whether from the first period to the sixty-fourth or from the first day to the sixteenth, or from the first two days to the last two days, the group exhibits marked improvement of performance, as measured in terms of either speed or accuracy of cancellation or both.

Section I. Transfer Effects in Terms of Speed and Accuracy of Performance.

The data to be discussed in Sections 1 and 2 will be found in Table VI, which shows the relation between the progress of the two groups. It is constructed as follows: In the first column the tests are listed; the numbers in the second column indicate the superiority of the practice group in gross speed, (all minus signs indicate inferiority of the practice group), in the third column the superiority in speed of cancellation is indicated, with the corresponding probable errors in the fourth column. The numbers in the fifth column indicate superiority in speed of cancellation in terms of per cent. of initial performance. The numbers in the sixth column represent superiority in per cent. of accuracy, with the corresponding probable errors in the seventh column. The eighth and ninth columns indicate superiority in performance when one is subtracted for each error (mostly omissions).

TABLE VI.

SHOWING RELATION BETWEEN PROGRESS OF PRACTICE GROUP AND CONTROL GROUP IN THE TEST SERIES

(The gains of the control group are subtracted from those of the practice group Minus signs indicate inferiority of the practice group.)

	SPEED				ACCURACY		SPEED PENALIZED FOR ERRORS	
	Gross	Net	P.E.	% Init. Perform.	%	P.E.	Net Efficiency	% Init. Perform.
Spanish a–t Words	6.74	6.44	0.38	67	1	1·92	6.12	74
Letters a and t	3.25	4.91	1.25	15	6	1.30	6.62	28
Capital A	4.53	3.99	1.39	10	–1	0.62	3.50	10
Capital B	2.13	1.13	0.93	5	–3	1.06	0.14	.7
Number Group 4 and 7	2.46	1.08	0.83	7	–7	1.13	–0.32	2
Letters e and s	–2.94	–3.89	1·78	–9	–3	0.75	–4.79	-13
Spanish e–s Words	3.32	0.02	0.36	2	–16	2.11	–3·28	–33

The table reads thus: In the Spanish a—t Word Test the practice group has a superiority in gross speed of 6.74 words per minute, and 6.44 cancellations per minute, with a probable error of 0.38, which in terms of initial performance is 67 per cent. In accuracy of performance the superiority is 1 per cent. with a probable error of 1.92. When one is subtracted for each error, the superiority is 6.12, which is 74 per cent. of the initial performance. The other tests are to be read in the same manner.

The results may be expressed in a different way as shown in Table VII, which is derived directly from Tables II and III. The table shows the percentage of the practice group reaching or exceeding the average performance of the control group in both speed and accuracy of cancellation. It reads thus: In the Spanish a—t Word Test 69 per cent. of the practice group reached or exceeded the average speed of the control group in the initial test, and 100 per cent. in the final test; while 66 per cent. reached or exceeded the average accuracy of the control group in the initial test, and 64 per cent. in the final test. The other tests are to be read in the same manner. The reader will

observe that the results as expressed in Table VII are in harmony with the results as expressed in Table VI. Consequently the discussion which follows will be based on the data as treated in Table VI.

TABLE VII.

SHOWING THE RELATION BETWEEN THE PRACTICE AND THE CONTROL GROUP IN TERMS OF THE PER CENT. OF THE PRACTICE GROUP REACHING OR EXCEEDING THE AVERAGE OF THE CONTROL GROUP

Tests	Percent of Practice Group reaching or exceeding Av. Speed of Control Group		Percent. of Practice Group reaching or exceeding Av. Per cent. of Accuracy of Control Group	
	Initial	Final	Initial	Final
Spanish a–t Words	69	100	66	64
Letters a and t	61	75	36	69
Capital A	69	75	77	72
Capital B	58	64	66	50
Number Group 4 and 7	50	55	80	33
Letters e and s	64	39	55	25
Spanish e-s Words	64	58	55	5

1. The first test is known as the Spanish a—t Word Test. It consisted of a paragraph of Spanish prose in which all words containing both a and t were to be cancelled. The stimuli determining the reaction in this test were not exactly the same as the stimuli determining the reaction in the practice series. The essential elements which had to be recognized and thus determine the cancelling were the same, but in Spanish words instead of English, and these in a Spanish context. In the initial test the practice group has the same speed of cancellation as in the first period of the practice series but a higher percentage of accuracy. However, in the final test neither the speed nor the accuracy of performance is as high as in the final practice period. This seems to indicate that in the beginning the cancelling in English prose was slightly more diffi-

cult than the cancelling in Spanish, while with prolonged practice the English a—t words become more familiar and accordingly more quickly cancelled. An examination of the practice material reveals the presence of many long and unusual a—t words. The fact that the subjects were very young and for the most part sons of foreign born parents explains their lack of familiarity with many of these English a—t words. Possibly then, the difference in accuracy may be explained by the fact that when one is cancelling prose in his own language, he tends at first to read the words, and just as his reading is impeded by the presence of unfamiliar words, so his accuracy of cancellation is interfered with. At any rate, the effects of practice in recognizing and cancelling a—t words in English was sufficient to give the practice group a decided superiority in the Spanish a—t Test. With an initial average performance of 10.03 cancellations per minute it has a superiority in improvement of 6.44 cancellations per minute, (P. E., 0.38) with neither superiority nor inferiority in the percentage of accuracy. The average initial performance of the two groups was 9.63 cancellations per minute. Therefore, the superiority of the practice group is 67 per cent. of the average initial performance of the two groups. When one considers that the practice group improved from the first period to the sixty-fourth in the practice series 159 per cent. in cancelling and 17 per cent. in accuracy; a superiority of 67 per cent. in cancelling is not so great in a test where the essential elements determining the cancelling were the same as those determining the cancelling in the practice series.

The results of this test are presented in graphic form in Figure V. The numbers along the vertical axis, both above and below the horizontal axis, represent the number of cases, while those along the horizontal *to the right* represent the *cancellation records*, and *to the left error records*. The graphs *above the horizontal* represent the performances of the *practice group*, while those below show the corresponding performances of the *control group*. The *initial performance* is indicated by *continous lines* and *final performance* by *dotted lines*. If one wishes to compare the speed of the practice group in the initial and final tests, he will compare the continuous curve with the dotted curve in the upper right quadrant; if one wishes to note the accuracy in the same way, he will compare the continuous curve with the dotted curve in the upper left

quadrant. If one wishes to note the performance of the control group, he will consult the lower right quadrant for change in speed and the lower left quadrant for change in accuracy.

The performances of each group in each of the other tests are shown in graphic form in Figures VI to XI, which are constructed in exactly the same manner as Figure V.

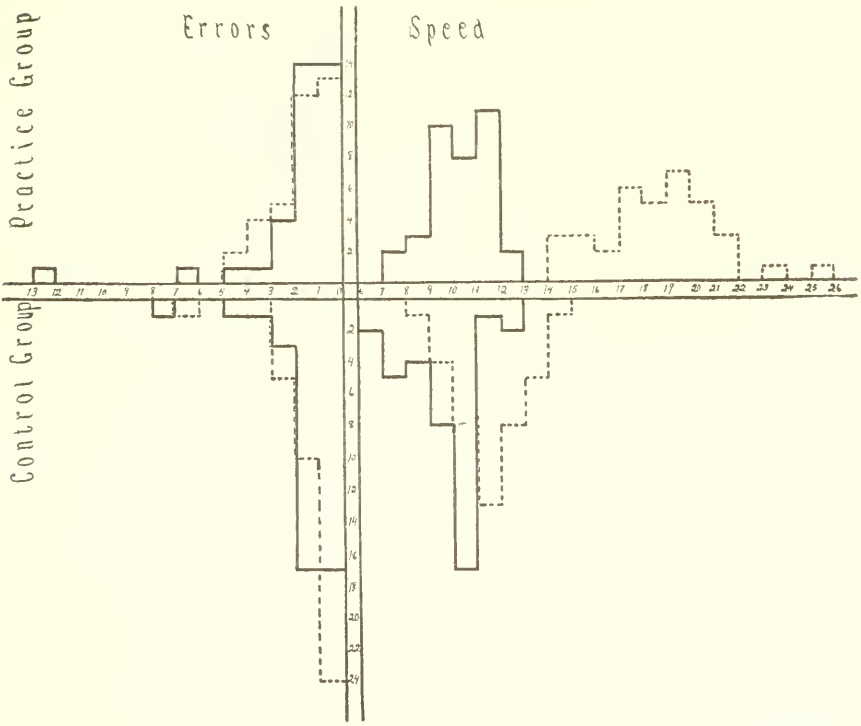

Figure V—Showing the Relation Between the Practice and Control Groups in the Spanish a—t Word Test.

2. The second test is known as the Spanish e—s Word Test. It consisted in cancelling in Spanish prose all words containing e and s. The essential elements which determined the cancelling in this test were among the elements to be neglected in the practice series, while the essential elements which determined the cancelling in the practice series were among the elements to be neglected in this test. In the practice only the a—t words were to be cancelled, while the e—s words along with all other words were to be neglected. In this test only e—s words were to be cancelled, while a—t words

along with all other words were to be neglected. With an initial average performance of 11.14 cancellations per minute, and an accuracy of 92 per cent., the practice group advanced to a final average performance of 12.54 cancellations per minute and decreased in accuracy to 76 per cent. With an initial average performance of 10.62 cancellations per minute, and an accuracy of 93 per cent., the control group advanced to a final average performance of 12.00 cancellations per minute, with no change in the per cent. of accuracy. In average performance as measured by the increase in the number of cancellations per minute there was no difference between the two groups. The practice was without effect so far as actual cancellations were concerned, for the two groups made the same improvement in speed of performance as measured by the increase in the number of cancellations per minute. In this test we see the importance of keeping the two characteristics of performance— speed and accuracy of cancellation—separate, for the transfer effects of practice are exhibited in the manifest inferiority of the practice group in accuracy of performance, which was 16 per cent. If we should arbitrarily correct for errors, the results would indicate an inferiority in speed of cancellation. But just how much inferiority in speed of cancellation corresponds to a loss of 16 per cent. in accuracy in this particular test no one knows at the present time. In this test it is probably large. However, the question arises: After prolonged practice in cancelling in English prose a—t words and neglecting e—s words along with all other words, could the group maintain its initial per cent. of accuracy when the task was cancelling e—s words in Spanish prose and neglecting a—t words along with all other words? The practice group had learned in actual experience the importance of accuracy as well as speed, and had acquired considerable accuracy as well as speed in cancelling a—t words in English, as revealed in its progress in the practice series. In only one test does it have a superiority in accuracy of performance, and that is the letter test, in which the letters a and t were to be cancelled. There is an interesting contrast between the performances of the practice group in the Spanish word tests. In the Spanish a—t Word Test it exhibits a superiority in speed of cancellation, while in the Spanish e—s Word Test it exhibits an inferiority in accuracy. In one test it gains as much in accuracy as the control group and has a decided superiority in speed of

cancellation; in the other it gains as much in speed of cancellation as the control group and has a decided inferiority in accuracy. In these two tests there is undoubted evidence of transfer. In the Spanish a—t Word Test the transfer effects are positive, appearing as facilitation. On the other hand, in the Spanish e—s Word Test the transfer effects are negative, appearing as interference. The conditions favoring facilitation are plainly evident. There was nothing in the test situa-

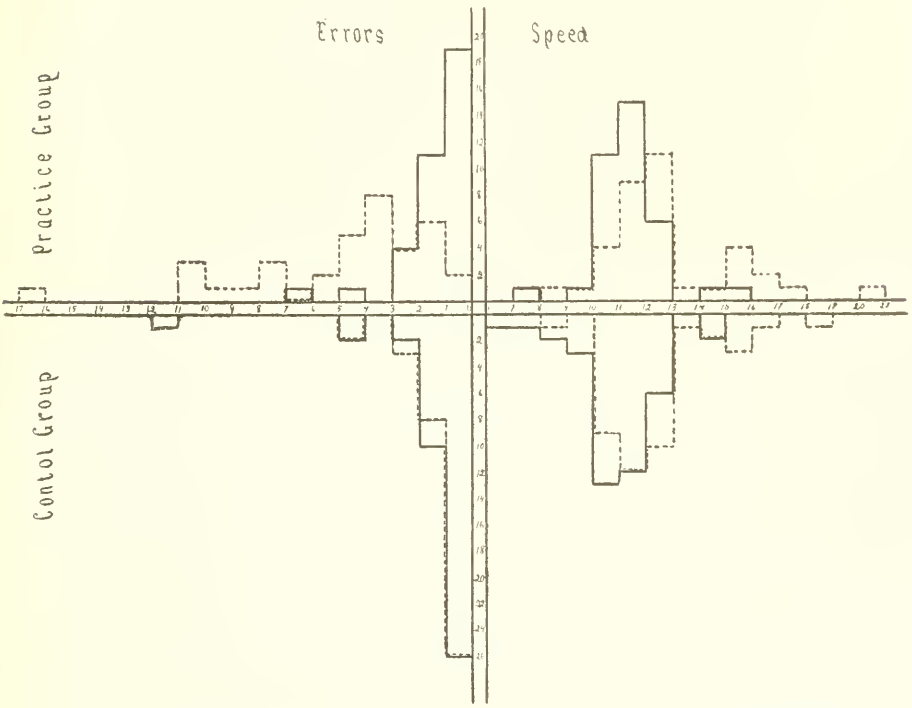

Figure VI—Showing the Relation Between the Practice and Control Groups in the Spanish e—s Word Test.

tion presented in the Spanish a—t Word Test to be cancelled which had to be neglected in the practice situation. The same elements in both situations determined the cancelling. Now, what were the conditions favoring interference? In what respects was the test situation presented in the Spanish e—s Word Test unlike the practice situation presented in the practice series? Were the two situations alike in any respects? The reader is already familiar with the practice situation pre-

sented in the practice series. The Spanish e—s Word Test situation was like it in that it contained many a—t words, and unlike it in being Spanish prose and in the task to be performed. The elements in the practice situation which the group had been trained to cancel were to be neglected; while some of the elements in the practice situation which it had been trained to neglect were now to be cancelled.

The graphic representation of the results of this test are shown in Figure VI, which is constructed in the same way as Figure V.

3. The third test is called the a and t Letter Test. The test situation consisted in a distribution of small letters, and the task consisted in cancelling the letters a and t. The practice group advanced from an average initial performance of 27.45 cancellations per minute and an accuracy of 87 per cent., to a final average performance of 37.35 cancellations per minute and an accuracy of 94 per cent. In the same test the control group advanced from an average initial performance of 26.00 cancellations per minute and an accuracy of 92 per cent., to a final average performance of 30.99 cancellations per minute and an accuracy of 93 per cent. The practice group gained 9.90 cancellations per minute with a probable error of 0.99, and in accuracy 7 per cent with a probable error of 1.02. The control group gained 4.99 cancellations per minute with a probable error of 0.77, and in accuracy 1 per cent. with a proable error of 0.81. Consequently, the practice group has a superiority of 4.91 cancellations per minute with a probable error of 1.25, and in accuracy 6 per cent. with a probable error of 1.30. The average of the initial performances in cancellation of the two groups is 26.72 per minute. In terms of this initial average, the superiority of the practice group is 15 per cent. Accordingly, the transfer effects of the practice are positive, appearing as facilitation in both speed of cancellation and accuracy of performance. There were no elements to be neglected in the test situation which were to be cancelled in the practice situation. In the practice situation the presence of both a and t in a word determined the cancelling, while in the test situation a and t themselves were to be cancelled. Thus, instead of cancelling a—t words the letters had to be cancelled. This required a slight change in the response. However, the differences were very slight compared with the similarities, for the essential elements determining the can-

celling were identical in both situations. In the next test the conditions are quite different and the results are in contrast with the results in the a and t Letter Test.

These results are shown graphically in Figure VII.

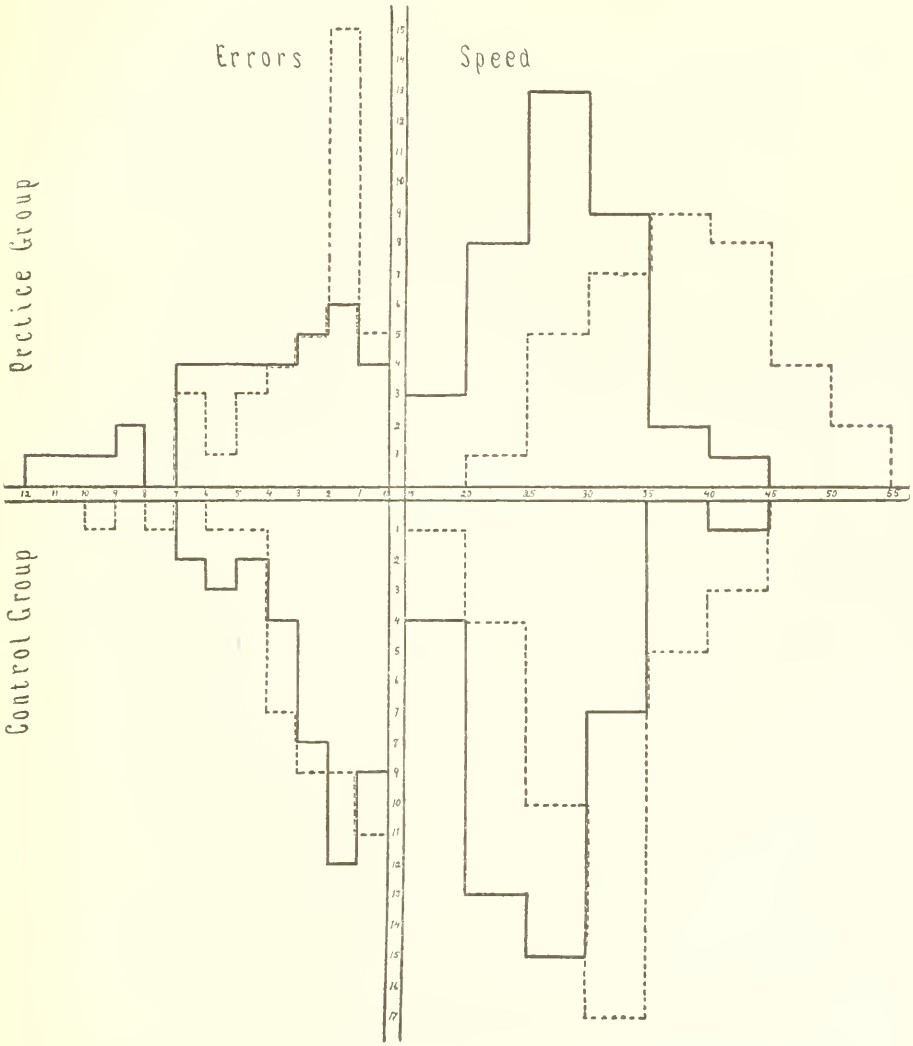

Figure VII—Showing the Relation Between the Practice and Control Groups in the a and t Letter Test.

4. The fourth test is called the e and s Letter Test. The test situation was the same as that in the third test—a distribution of small letters—while the task consisted in cancelling the letters e and s, instead of the letters a and t. The practice group advanced from an average initial performance of 41.26 cancellations per minute to a final average performance of 43.51 cancellations per minute, and decreased in accuracy from 95 per cent. to 94 per cent. In the same test the control group advanced from an average initial performance of 39.00 cancellations per minute and an accuracy of 96 per cent., to a final average performance of 45.14 cancellations per minute and an accuracy of 98 per cent. The practice group gained 2.25 cancellations per minute with a probable error of 1.29, and decreased in average accuracy 1 per cent. with a probable error of 0.66. The control group gained 6.14 cancellations per minute with a probable error of 1.23 and in accuracy 2 per cent. with a probable error of 0.35. Consequently, the practice group has an inferiority of 3.89 cancellations per minute with a probable error of 1.78 and in average accuracy an inferiority of 3 per cent. with a probable error of 0.75. The initial average of the two groups is 40.40 cancellations per minute. In terms of this initial average, the inferiority of the practice group in speed of cancellation is 9 per cent. The transfer effects of the practice are slightly negative, appearing as interference both in speed of cancellation and in accuracy of performance. The results in this test are in contrast with the results in the a and t letter test. In the two tests the objective situations were identical, but the tasks to be performed different. In the one it was the a and t aspect that caused the cancelling, in the other the e and s aspect. But the a and t aspect as a cause of cancelling had been trained in the practice series, while the e and s aspect had been neglected. The attempt to neglect the a and t aspect and respond to the e and s aspect resulted in interference. The results in the two pairs of tests, the Spanish a—t and e—s Word Tests and the Letter Tests a and t and e and s, indicate clearly the conditions favoring facilitation on the one hand and interference on the other.

The results of this test are represented graphically in Figure VIII.

5. The fifth test is called the Columbia A Test. It consisted in a distribution of capital letters, and the task to be performed consisted in cancelling capital A's. In this test the

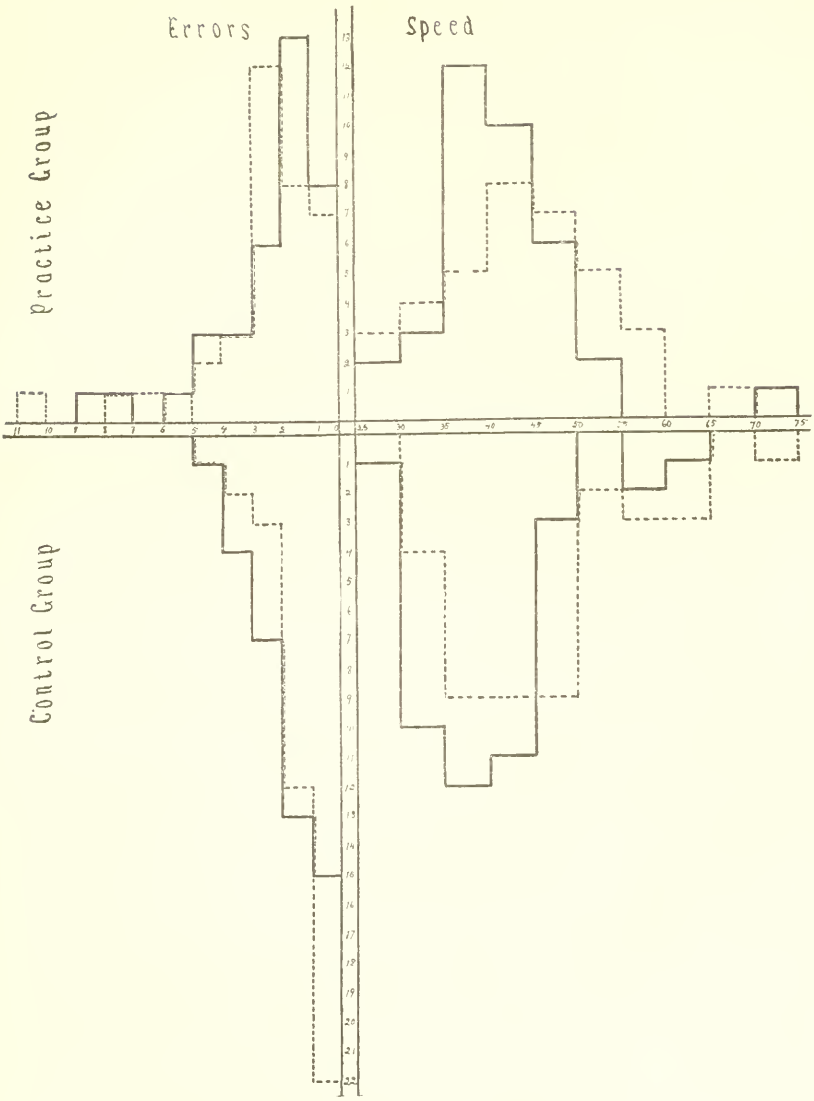

Figure VIII—Showing the Relation Between the Practice and Control Groups in the e and s Letter Test.

practice group advanced from an initial average performance of 38.33 cancellations per minute and an accuracy of 97 per cent., to a final average performance of 51.75 cancellations per minute and an accuracy of 98 per cent. In the same test the control group advanced from an initial average performance of

36.12 cancellations per minute and an accuracy of 97 per cent., to a final average performance of 45.55 cancellations per minute and an accuracy of 99 per cent. The practice group gained 13.42 cancellations per minute (P. E. 1.07) and in accuracy 1 per cent. (P. E. 0.46). The control group gained 9.43 cancellations per minute (P. E. 0.88) and in accuracy 2 per cent. (P. E. 0.42). The practice group has a superiority of 3.99 cancellations per minute, (P. E. 1.39), and in accuracy an inferiority of 1 per cent. (P. E. 0.62). In this test the transfer effects are positive, causing a very slight facilitation in the speed of cancellation. The initial average performance of the two groups was 37.22 cancellations per minute. In terms of this average the superiority of the practice group in speed of cancellation is 10 per cent. What were the conditions favoring transfer? An examination of the practice situation reveals very few capital A's. These few were to be neglected only when they were in words not containing t. Their appearance was so rare that they seldom determined the cancelling or had to be neglected. It is, consequently, fair to assume that capital A figured rarely in the practice series. If so, then, why this slight transfer effect in the capital A Test? Identity in name with one of the letters determining the cancelling in the practice situation.

The graphic representation of the results of this test is shown in Figure IX.

6. This test is known as the Thorndike B Test. It consisted in a distribution of capital letters, and the task consisted in cancelling capital B's. The practice group advanced from an average initial performance of 20.93 cancellations per minute to a final average performance of 26.00 cancellations per minute, and decreased in accuracy from 96 per cent. to 93 per cent. The control group advanced from an initial average performance of 19.81 cancellations per minute and an accuracy of 96 per cent. to a final average performance of 23.75 cancellations per minute, with no change in accuracy. The practice group gained in average performance 5.07 cancellations per minute (P. E. 0.71), and lost in average accuracy 3 per cent. (P. E. 0.87). The control group gained in average performance 3.94 cancellations per minute (P. E. 0.60), with no change in accuracy. The practice group has a superiority of 1.13 cancellations per minute (P. E. 0.93), and an inferiority in accuracy of 3 per cent. (P. E. 1.06). It is evident that the

transfer effects in speed of cancellation are not positive, for 1.13 (P. E. 0.93) is unreliable. However, 3 per cent. (P. E. 1.06) is fairly reliable and indicative of interference. It may be of interest to contrast the capital B Test with the capital A Test. In the A Test the practice group has a superiority in speed of 3.99 cancellaions per minute with no reliable change

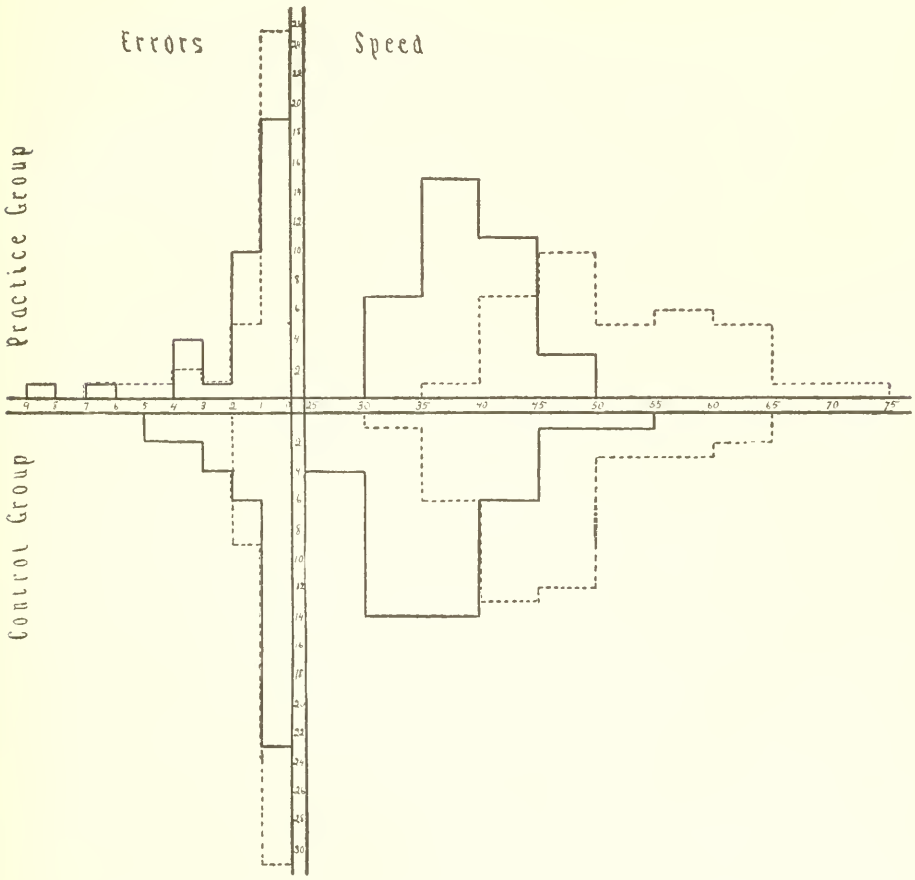

Figure IX—Showing the Relation Between the Practice and Control Groups in the Columbia A Test.

in accuracy. In the B Test there is no reliable evidence of transfer in speed of cancellation, but fairly reliable evidence of loss in accuracy. The question arises; What caused the transfer effects to appear in the A Test as a slight facilitation and in the B Test as a slight interference? An analysis of the two

tests in their relation to the practice situation reveals the conditions favoring a slight transfer in the one positively, in the other negatively. In the A Test there were 100 capital A's distributed among other capital letters, and among these 15 capital T's. In the B Test there were 50 capital B's distributed among other capital letters, and among these there were 50 capital A's and 25 capital T's. In the practice situation, there were comparatively few capital B's and numerous small b's These were to be neglected whether the same in name or in form. In the practice situation there were comparatively few capital A's and numerous small a's. These, whether the same in form or name, were to be neglected except when in words containing t also. When they appeared with t in words, such words were to be cancelled. A test situation in which all the elements to be neglected were elements to be neglected in the practice situation with the exception of a few capital T's, and in which the element to be cancelled was the same in name as one of the elements to be recognized in the practice situation and to determine cancelling when it appeared with t in a word, should be more favorable to positive transfer than a test situation in which a large number of the elements to be neglected—50 capital A's and 25 capital T's—were the elements, identical in name and rarely in form, with the elements to be recognized in the practice situation and to determine the cancelling, and further, in which test situation the element determining the cancelling was identical in name and rarely in form with one of the elements to be neglected in the practice situation. Thus it is seen that conditions in the A Test situation slightly favored facilitation while conditions in the B Test situation slightly favored interference.

Figure X is a graphic representation of these results.

7. The seventh and last test is known as the Woodworth and Wells Number Group Checking Test. The task consisted in cancelling six-place number-groups containing both 4 and 7. The practice group advanced from an average initial performance of 15.51 cancellations per minute to a final average performance of 20.47 cancellations per minute, and decreased in accuracy from 97 per cent. to 92 per cent. The control group advanced from an average initial performance of 14.87 cancellations per minute and an accuracy of 96 per cent. to a final average performance of 18.75 cancellations per minute and an accuracy of 98 per cent. The practice group gained 4.96 can-

cellations per minute (P. E. 0.70) and decreased in accuracy 5 per cent (P. E. 0.97). The control group gained 3.88 cancellations per minute (P. E. 0.44), and 2 per cent. in accuracy (P. E. 0.58). The practice group has a superiority in speed of

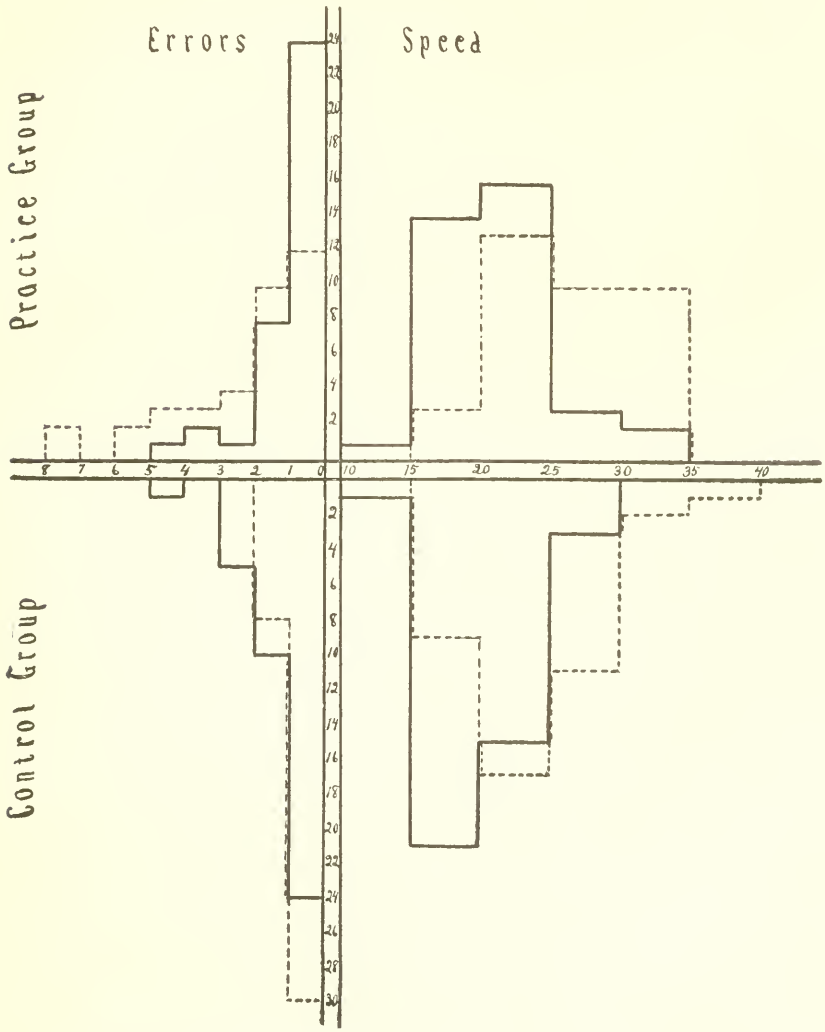

Figure X—Showing the Relation Between the Practice and Control Groups in the Thorndike B Test.

1.08 cancellations per minute (P. E. 0.83) and an inferiority in accuracy of 7 per cent. (P. E. 1.13). There is no reliable evidence of transfer in speed of performance, for 1.08 with a P. E. of 0.83 is unreliable. But an inferiority in accuracy of

7 per cent. (P. E. 1.13) indicates transfer effects in the form of interference.

What were the relations between the practice situation and the situation presented i nthe Number-Group Checking Test that favored interference? None of the essential elements determining the cancelling in the practice situation were present in the test situation. In fact, so far as content elements are concerned, the two situations had nothing in common. Letters in letter groups made up the content of the practice situation, while the content of the test situation consisted of numbers in number groups. It was both a and t in a letter group that determined the cancelling in the practice situation, while the cancelling in the test situation was determined by the presence of both 4 and 7 in a number group. While the presence of both a and t in a letter group determined the cancelling, yet, it was the letter group that was to be cancelled and not the letters. So with the presence of 4 and 7. Not 4 or 7, but the number group containing both 4 and 7 was to be cancelled. Consequently, the two situations were entirely dissimilar as to content elements, but somewhat similar and suggestive of likeness in the fact that the content elements in both situations were in group forms, and while different content elements determined the cancelling, yet it was in both situations the group form that was to be cancelled.

Figure XI represents these results graphically.

Section 2. Transfer Effects in Terms of Gross Speed; and When Arbitrarily Corrected for Errors.

The results will now be presented and discussed as they appear when 1 is subtracted for each error, and also, when errors and cancellations are added, to show ground covered or gross speed. The writer has previously stated his objections to penalizing for errors. However, a presentation of the results of such arbitrary procedure will indicate the grounds of this objection. In the Spanish a—t Word Test, correcting for errors gives the practice group a superiority of 6.12 cancellations per minute instead of 6.44. Here, correcting makes little difference, as each group gains about the same amount in percent of accuracy. In gross speed the superiority of the practice group is 6.74 words per minute. However, in the Spanish e—s Word Test correcting for errors does make a difference, for it gives the practice group an inferiority of 3.28 cancella-

tions per minute, when as a matter of fact, both groups made equal gains in speed of cancellation, while the practice group decreased 16 percent in accuracy. Accepting the corrected results would lead to the inference that the transfer effects

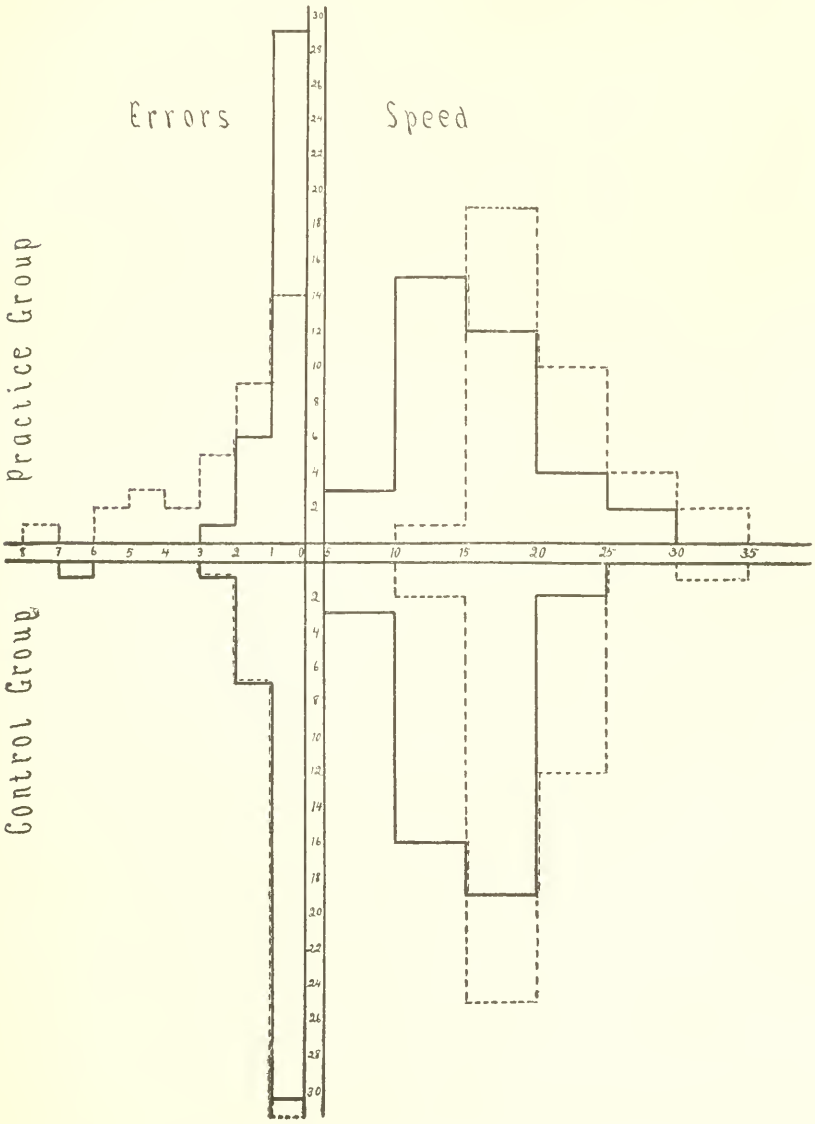

Figure XI—Showing the Relation Between the Practice and Control Groups in the Woodworth and Wells Number-Group Checking Test.

of the practice caused an inferiority in speed of cancellation. In gross speed the practice group has a superiority of 3.32 words per minute. What, then, were the transfer effects in this test? Certainly not an inferiority in speed of cancellation, as indicated by the corrected results. Taking the facts as they actually occur, we note that the training of the practice series caused a superiority in gross speed, and an inferiority in percent of accuracy, and no difference in speed of cancellation. Surely, no one would say that a gross speed of 3.32 words per minute and an inferiority in accuracy of 16 percent with no difference in speed of cancellation, means the same thing as an inferiority of 3.28 cancellations per minute.

In the a and t Letter Test, corrected results give the practice group a superiority of 6.62 cancellations per minute, instead of 4.91 with a superiority of 6 percent in accuracy, and in gross speed of 3.25 words per minute. Thus, in this test, as in the Spanish e—s Word Test, corrected results tend to conceal the actual transfer effects of the practice, for instead of a superiority in cancellation only, the practice group has a superiority in gross speed and in speed and accuracy of cancellation. In the e and s Letter Test, corrected results give the practice group an inferiority of 4.79 cancellations per minute, while the actual results indicate that the practice caused an inferiority, not only in the speed of cancellation, but in accuracy of performance and in gross speed. The transfer effects in these two tests are in contrast whether the results are corrected or not. However, it is plainly evident that the uncorrected results give a clearer and truer presentation of the transfer effects than the corrected results, for instead of the practice causing an inferiority of 4.79 cancellations per minute, it caused an inferiority in gross speed of 2.94 words per minute, in speed of cancellation 3.98 per minute, and in accuracy of performance 3 percent.

Corrected results in the capital A Test indicate that the practice group has a superiority of 3.5 cancellations per minute, while with the uncorrected results it has a superiority of 3.99. Since there is no reliable difference in the gains in accuracy made by the two groups, the transfer effects are manifested in speed of cancellation, and in gross speed which was 4.53 words per minute. In the capital B test, corrected results indicate a superiority of 0.14 cancellations per minute

for the practice group, while uncorrected results give an actual though unreliable superiority of 1.13 cancellations, and a reliable inferiority in accuracy of 3 percent. In these two tests corrected results indicate positive transfer effects in the Capital A Test, and none at all in the Capital B Test. On the other hand, the actual results indicate transfer effects in the Capital A Test in a superiority of actual cancellations, and in gross speed; and in the Capital B Test an inferiority in accuracy of 3 percent., and a superiority in gross speed of 2.13 words per minute.

In the Number Group-Checking Test corrected results indicate an unreliable inferiority of 0.32 cancellations per minute, on the part of the practice group, while the actual results show a reliable inferiority in accuracy of 7 percent, and in gross speed a superiority of 2.46 words per minute. The writer feels confident that this presentation and discussion of the results of correcting for errors justifies his objection to a procedure so arbitrary; a procedure which tends to conceal the actual transfer effects of the practice.

Section 3. Summary and Conclusion.

The transfer effects of the practice are revealed in the superiority or inferiority of the practice group. These effects may manifest themselves in three ways: In gross speed, that is, total ground covered cancellations plus errors; in speed of actual cancellation; and in accuracy of performance. Prolonged practice in cancelling in English prose words containing both a and t resulted in an improvement in gross speed from an average of 129.6 words per ten minute period to an average of 277.0, or a gain of 113 percent; in speed of cancellation from an average of 102.6 cancellations per period to an average of 266.5, or a gain of 159 percent; and in accuracy of cancellation from 79 percent to 96 percent, or a gain of 17 percent. In gross speed the practice group has a superiority in six of the seven tests and an inferiority in one. In speed of cancellation the transfer effects that are reliable are positive in three tests; being 67 percent in the Spanish a—t Word Test, 15 percent in the a and t Letter Test, and 10 percent in the capital A Test; and negative in one, being 9 percent in the e and s Letter Test. In percent of accuracy the transfer effects are positive in one test and negative in four tests, and without

reliable influence in the other two. In only one test does the practice group have a superiority in both speed and accuracy of cancellation—the a and t Letter Test; and in only one test is it inferior in these two respects—the e and s Letter Test. In each of the seven tests there is evidence of the transfer effects of the practice, and these are pronounced in four tests; in two there is facilitation and in two there is interference. The two Spanish Word Test situations were identical; in the one the stimulus provoking cancellation was identical in form and content elements with that in the practice situation. These were the conditions favoring the largest amount of positive transfer or facilitation. In the other, the stimulus provoking cancellation was one of the stimuli to be neglected in the practice situation, while the practiced stimulus was neglected. Such conditions favored interference. The two Letter Test situations were identical; in the one the stimuli provoking cancellation were identical with the content elements of the stimulus cancelled in the practice situation. Such conditions favored facilitation. In the other, the stimuli provoking cancellation were unlike the content elements of the stimulus cancelled in the practice, and like the content elements of many stimuli neglected in the practice. Such conditions favored interference. In the two Capital Letter Tests the situations were somewhat unlike and the tasks different. In the Capital A Test the stimulus provoking cancellation was the same in name as one of the stimuli determining the cancelling in the practice situation. The result was a slight facilitation in the form of speed of cancellation. In the Capital B Test, the stimulus provoking cancellation was the same in name and rarely in form as one of the stimuli neglected in the practice situation, while there were many stimuli to be neglected that were the same in name and rarely in form as the stimuli determining cancelling in the practice situation. The result was a slight interference in the form of a decrease in accuracy. In the Number Group-Checking Test the stimulus provoking the cancellation response was the presence of 4 and 7 in a number group, while in the practice, it was the presence of a and t in a letter group. The result was interference in the form of a decrease in accuracy.

The tendency toward speed acquired through the practice was stronger than the tendency toward accuracy, for in six

of the seven tests, the practice group has a superiority in gross speed. However, in only three of the six, does this gross speed eventuate in facilitation; in the other three it is accompanied by interference in the form of a decrease in accuracy; and in the remaining one was not sufficient to prevent a loss in gross speed and in speed of cancellation, but was sufficient to cause a decrease in accuracy.

Section 4. Generalized Conclusions.

In this investigation the writer endeavored to discover under what conditions, to what extent, and in what direction, prolonged cancelling in one situation affects cancelling in other situations. The results justify the following conclusions:

1. That the extent and direction of the transfer effects depend, first, on the practice situation and the task performed; and second, on the test situation and the task performed.

2. When the test situation is such that the task performed is identical with the task in the practice, the extent of the transfer effects is largest and in a positive direction in the form of speed.

3. When the test situation is such that the task performed is the reverse of that performed in the practice, the extent of the transfer effects is large, causing **interference** in the form of a decrease in accuracy.

4. When the test situation is such that the task performed is identical as to content elements, with that of the practice, the extent of the transfer effects is comparatively small and in a positive direction, both in speed and accuracy.

5. When the test situation is such that the task performed is the reverse as to content elements of that of the practice, the transfer effects are small causing interference both in speed and accuracy.

6. When the test situation is such that the task performed is somewhat similar as to content elements to that of the practice, the transfer effects are slightly positive in the form of speed.

7. When the test situation is such that the task performed is somewhat the reverse as to content elements of that in the practice, the transfer effects are slightly negative in the form of a decrease in accuracy.

8. When the test situation contains none of the content elements of the practice situation but is such that the task performed is somewhat similar as to form with that of the practice, the transfer effects are negative, causing a decrease in accuracy.

9. Finally, this investigation plainly teaches the importance of scoring for accuracy as well as speed.

VITA

Melvin Albert Martin. Born at Petersburg, Virginia, October 18, 1871. Early education received in the Public Schools of Petersburg and at the Virginia Polytechnic Institute. A.B., Richmond College, Richmond, Virginia, 1898; A.M., Columbia University, 1905. Special student in the University of Chicago, 1898-1900; Head Master, Mossy Creek Academy, Virginia, 1900-1901; Principal Southside Female Institute, Burkeville, Virginia, 1901-1902; professor of Philosophy and Physics, Woman's College, Richmond, Virginia, 1902-1904; graduate student, Columbia University, 1904-1905; Dean and professor of Philosophy and Education, Woman's College, Richmond, Virginia, 1905-1913; professor of Psychology, in the Summer School of the University of Virginia, 1911-14; graduate student, Columbia University, 1913-1915; assistant in Psychology, Columbia University, 1914-1915.

LIBRARY OF CONGRESS
0 021 350 395 9